

## SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

### Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

### STAFF USE ONLY

Date completed: 05-28-03  
Searcher: Berc. Uy 24954  
Terminal time: 20  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: 25  
Number of Searches: \_\_\_\_\_  
Number of Databases: 1

#### Search Site

\_\_\_\_\_ STIC  
\_\_\_\_\_ CM-1  
\_\_\_\_\_ Pre-S

#### Type of Search

\_\_\_\_\_ N.A. Sequence  
\_\_\_\_\_ A.A. Sequence  
\_\_\_\_\_ Structure  
\_\_\_\_\_ Bibliographic

#### Vendors

\_\_\_\_\_ IG Suite  
\_\_\_\_\_ STN  
\_\_\_\_\_ Dialog  
\_\_\_\_\_ APS  
\_\_\_\_\_ Geninfo  
\_\_\_\_\_ SDC  
\_\_\_\_\_ DARC/Questel  
\_\_\_\_\_ Other CGN

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 22:35:43 ; Search time 3485 Seconds  
(without alignments)  
3006.316 Million cell updates/sec

Title: US-10-023-586B-2

Perfect score: 1936

Sequence: 1 MSLILPSRSGSRGAL.....QNKLGHPAGKRCPCGLNRS 360

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q/cyg2\_1/USPTO\_spool/US10023586/runat\_22052003\_114850\_28641/app\_query.fasta\_1.519  
-DB=GenEmbl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pcpt -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10023586.CGN\_1.1.2496.erunat\_22052003\_114850\_28641 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pin.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1936	100.0	1082	6	AX458306	Sequence
2	1936	100.0	1334	9	AB065670	Homo sapi
3	1936	100.0	3143	6	AX365511	Sequence
c 4	1936	100.0	164502	2	AC026960	Homo sapi
5	1936	100.0	167084	9	AC092999	Homo sapi
6	1923	99.3	163958	2	AC021773	Homo sapi
7	1806	93.3	1019	9	AF411108	Homo sapi
8	1785	92.2	1020	6	AX458308	Sequence
9	1785	92.2	1076	6	AX375230	Sequence
10	1541	79.6	850	6	AX365514	Sequence
11	1154.5	59.6	209273	2	AC121579	Mus muscu
12	1130	58.4	91107	2	AC116265	Rattus no
c 13	921	47.6	510	6	AX147768	Sequence
c 14	888	45.9	192524	2	AC120742	Rattus no
c 15	732	37.8	681	6	AX244718	Sequence
16	681	35.2	585	6	AX244727	Sequence
17	501	25.9	1074	5	AF069555	Meleagris
18	501	25.9	1977	5	AF031897	Meleagris
19	495	25.6	1163	5	GDP2Y3	G.domesticu
20	460.5	23.8	1086	10	MM0277752	Mus muscu
21	460.5	23.8	215102	2	AL671299	Mus muscu
22	460.5	23.8	215410	10	AC091784	Genomic s
23	453	23.4	123185	2	AP000587	Homo sapi
24	447	23.1	2681	9	BC012104	Homo sapi
25	446.5	23.1	1651	9	HSU40223	Human uridi
26	446.5	23.1	136930	5	AL590151	Zebrafish
27	446	23.0	1177	10	RNP2Y4REC	R.norvegicu
28	445.5	23.0	1232	9	HSGPCRE	H.sapiens g
29	445.5	23.0	1429	6	A62985	Sequence 1
30	445.5	23.0	1429	6	HSP2Y4	Sequence 1
c 31	445.5	23.0	178861	2	AC068231	Homo sapi
32	445.5	23.0	178868	9	AL357752	Human DNA
33	442	22.8	2025	9	HSU07225	Homo sapien
34	442	22.8	2257	9	BC028135	Homo sapi
35	442	22.8	2684	10	RNP2Y4RG	Rattus norv
36	436	22.5	1922	10	RATNGPCF2R	Rat mRNA fo
37	435	22.5	196988	9	AF002761	Homo sapi
38	434.5	22.4	2361	5	XP2Y8	X.laavis mR
39	420.5	21.7	984	6	AR109625	Sequence
40	420.5	21.7	984	6	E12485	Mouse cDNA
41	420.5	21.7	987	10	AF298899	Mus muscu
42	420.5	21.7	1020	6	AR109898	Sequence
43	420.5	21.7	1976	10	BC027331	Mus muscu
44	420	21.7	1842	6	I34573	Sequence 1
45	420	21.7	1842	6	I36524	Sequence 1

ALIGNMENTS

RESULT 1

```
AX458306
LOCUS AX458306 1082 bp DNA linear PAT 08-JUL-2002
DEFINITION Sequence 1 from Patent EP1215214.
ACCESSION AX458306
VERSION AX458306.1 GI:21725001
KEYWORDS human.
SOURCE Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1
REFERENCE
  Fidock, M.D.
  Novel polypeptide
  Patent: EP 1215214-A 1 19-JUN-2002;
  Pfizer Limited (GB) ; PFIZER INC. (US)
FEATURES
  Location/Qualifiers
    1..1082
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
BASE COUNT 218 a 318 c 277 g 269 t
ORIGIN
Alignment Scores:
  Pred. No.: 3 68e-166 Length: 1082
  Score: 1936.00 Matches: 360
  Percent Similarity: 100.00% Conservative: 0
  Query Local Similarity: 100.00% Mismatches: 0
  Query Match: 100.00% Indels: 0
  DB: 6 Gaps: 0
US-10-023-586B-2 (1-360) x AX458306 (1-1082)
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20
  |||||
DB 1 ATGCTGTCCATTGCTTCTCCAGGGGAAGCAAGAGCGGCGCTGCTGGAGCTCG 60
QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
  |||||
DB 61 CTCCTGGAGGAGCCCTCCCGGACATGGAGAGGTGGACATGAATACATCAGGAACAA 120
QY 41 GlyLeuCysGlnPheSerGlnLysTyrTyrGlnValTyrLeuSerLeuAlaTyrSerIle 60
  |||||
DB 121 GGTCTCTGCAGTCTTCACAGAAAGTACAAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 180
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheIleSerTrpGlyGlnThr 80
  |||||
DB 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGGCCAAACC 240
QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
  |||||
DB 241 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCCGACCTGCTTTAT 300
QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspAspArgTrpProPheGlyGlu 120
  |||||
  301 GTGCTATTGCCCTTCTCATCATCACTACTACTACATGACAGGTGGCCCTTCGGGGAG 360
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
  |||||
  361 CTGCTCTGCAAGCTGGTGCACCTCTCTGTATATCAACCTTTAGCGGACGATCTGCTG 420
DB 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuSerLeuPro 160
  |||||
DB 421 CTGACCTGATCTCTGTGCACCACTCTCTAGGTGTGTGGCCACCACTGTGTTCGTGGCC 480
QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180
  |||||
DB 481 TACCGGACCGCAGCATCCCTGGCTGGCGACACCACTACCACTGGGCCCTGTGTCTCTC 540
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
  |||||
DB 541 CAGCTGCTGCCACACTGGCTCTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 600
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
  |||||
```

```
|||||
DB 601 TATGACATGACAGCCCAAGAGAATTTTGATCGCTTTTTCCTACGCGATAGTCTTGACA 660
QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
  |||||
DB 661 TTGCTCGGCTTTCTTTCCCTCTGCTGCTATTTTGGTGTCTATTTCACTGATGCTCAGGAG 720
QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
  |||||
DB 721 CCTGATCAGCCAGAGGAGAACCTCATGAGGACAGGCAACACAGCCGAGCCAGGTCAT 780
QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
  |||||
DB 781 CGGGACCATCTACTGCTGTGGCTCTTCACTCTCTCTTTTGTGCCCTTCCATATCAC 340
QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
  |||||
DB 841 TCGCTCTCTTCTACCTACCATCTGCTTTCTCTGAGGACTGCGAGCTCTTGATGCG 900
QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
  |||||
DB 901 AGCCAGTGTGGCTACAAGATATGGAGGCTCTGCTGAGTGTGAGCAGCTGCTCAACC 960
QY 321 SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340
  |||||
DB 961 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCTCAGAACTGAGG 1020
QY 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
  |||||
DB 1021 CAGACAACTGGGTGAGCCTCAGCTGGGAGGAAGAGATGCCAGGGTGAACAGATCT 1080
RESULT 2
LOCUS AB065670 1334 bp DNA linear PRI 23-JUL-2002
DEFINITION Homo sapiens gene for seven transmembrane helix receptor,
  isolate:CBRC7TM_233.
ACCESSION AB065670
VERSION AB065670.1 GI:21928614
KEYWORDS
SOURCE Homo sapiens (isolate:CBRC7TM_233) DNA.
ORGANISM Homo sapiens
  Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1
REFERENCE
  Suwa, M., Sato, T., Okouchi, I., Arita, M., Futami, K., Matsumoto, S.,
  Tsutsumi, S., Aburatani, H., Asai, K. and Akiyama, Y.
  Genome-wide discovery and analysis of human seven transmembrane
  helix receptor genes
  Unpublished
  2 (bases 1 to 1334)
  Direct Submission
  Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research
  Center (CBRC), National Institute of Advanced Industrial Science
  and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan
  (E-mail: m-suwa@aist.go.jp, URL: http://www.cbrc.jp/,
  Tel: 81-3-3599-8080, Fax: 81-3-3599-8081)
  This sequence is a seven transmembrane helix receptor candidate
  system that contains programs of gene
  finding (GeneDecoder), sequence search, motif-domain assignment and
  transmembrane helix prediction.
  And the sequence is submitted by the collaborative project between
  [Computational Biology Research Center (CBRC), National Institute
  of Advanced Industrial Science and Technology (AIST)] and [Genome
  Science Division, Research Center for Advanced Science and
  Technology (RCAT), University of Tokyo].
  Location/Qualifiers
    1..1334
     /organism="Homo sapiens"
     /isolate="CBRC7TM_233"
     /db_xref="taxon:9606"
     /chromosome="3"
FEATURES
  source
```







\* 36238 36337: gap of 100 bp  
\* 36338 46377: contig of 10040 bp in length  
\* 46378 46477: gap of 100 bp  
\* 46478 59361: contig of 12884 bp in length  
\* 59362 59461: gap of 100 bp  
\* 59462 75221: contig of 15760 bp in length  
\* 75222 75321: gap of 100 bp  
\* 75322 93371: contig of 18050 bp in length  
\* 93372 93471: gap of 100 bp  
\* 93472 117463: contig of 23992 bp in length  
\* 117464 117563: gap of 100 bp  
\* 117564 137715: contig of 20152 bp in length  
\* 137716 137815: gap of 100 bp  
\* 137816 164502: contig of 26687 bp in length.

FEATURES

source  
1. 164502  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3"  
/clone="RP11-170K4"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. .2917  
/note="assembly\_fragment"  
3018. 6310  
/note="assembly\_fragment"  
6411. 9447  
/note="assembly\_fragment"  
9548. 14470  
/note="assembly\_fragment"  
14571. 20528  
/note="assembly\_fragment"  
20629. .24667  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:right"  
24768. 29367  
/note="assembly\_fragment"  
29468. .36237  
/note="assembly\_fragment"  
36338. 46377  
/note="assembly\_fragment"  
46478. 59361  
/note="assembly\_fragment"  
59462. .75221  
/note="assembly\_fragment"  
75322. .93371  
/note="assembly\_fragment"  
93472. 117463  
/note="assembly\_fragment"  
117564. .137715  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:left"  
137816. .164502  
/note="assembly\_fragment"  
BASE COUNT 48539 a 32273 c 33137 g 49088 t 1405 others  
ORIGIN

Alignment Scores:  
Pred. No.: 2.7e-163 Length: 164502  
Score: 1936.00 Matches: 360  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-10-023-586b-2 (1-360) x AC026960 (1-164502)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
|||||  
Db 123728 ATGCTGCCATTTGCTTCTCCAGGGAAGCAGAGCGGAGCGCTGCTGGAGCTCTG 123669

Qy 21 LeuLeuGluGluAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
|||||  
Db 123668 CTCTGGAGGAGCCCTCCCGGACATGGAGAGGTGGACATGAATACATCACAGGAACA 123609

Qy 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60  
|||||  
Db 123608 GGTCCTGCCAGTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCTACAGATATC 123549

Qy 61 IlePheIleLeuGluLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80  
|||||  
Db 123548 ATCTTTATCTAGGGCTGCCACCTAAATGGCACCTGTCTGTGGCACCTCTGGGGCAAC 123489

Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
|||||  
Db 123488 AAGCGTGGAGCTGTGCCACCACTATCTGGTGAACCTGTATGGTGGCGACCTGCTTTAT 123429

Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpProPheGlyGlu 120  
|||||  
Db 123428 GTGCTATTGCCCTTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAG 123369

Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
|||||  
Db 123368 CTGCTCTGCAAGCTGGTGCACCTTCTCTGTATATCAACCTTTACGGCAGCATCTGCTG 123309

Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
|||||  
Db 123308 CTGACCTGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTGTCCGCTGCC 123249

Qy 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValLeu 180  
|||||  
Db 123248 TACCGACCCGAGGATGCTGGCTGGCCAGCACCACTGGCCCTGGTGGTCTC 123189

Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200  
|||||  
Db 123188 CAGCTGCTGCCACACTGGCTTCTCCACACGAGTACATCAATGCCAGATGATCTGG 123129

Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220  
|||||  
Db 123128 TATGACATGACCAAGCAAGAGAAATTTGATCGGCTTTTGGCTACGCGCATAGTTCTGACA 123069

Qy 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
|||||  
Db 123068 TTGCTGGCTTCTTCCCTTCTGGTCTATTTGGTGTGCTATCTACTGATGGTCAAGAG 123009

Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260  
|||||  
Db 123008 CCTGATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCGAGCCAGGTCCAT 122949

Qy 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280  
|||||  
Db 122948 CCGGACCATCTACTGTGTGGCTCTTCAACCTCTGTGTGGCTTCCATATCAC 122889

Qy 281 SerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300  
|||||  
Db 122888 TCGCTCTTCTACCTCACCATCTGCTTCTTCTCAGGACTGCCAGCTCTTGATGGC 122829

Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnGlnLeuProGlnPro 320  
|||||  
Db 122828 AGCCAGTGTGGCTTACAAGATATGGAGGCTCTGGTGGTGTGAGCAGCTGCCCAACCC 122769

Qy 321 SerProValLeuSerPheLeuGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340  
|||||  
Db 122768 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAAGGTCTCTCCAGAACTGAGG 122709

Qy 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
|||||  
Db 122708 CAGAACAAAGTTGGTGGATCTCCAGCTGGGAGGAGAGATGCCCGAGGTTCACACATCT 122649

RESULT 5  
AC092999  
LOCUS  
DEFINITION

AC092999 167084 bp DNA linear PRI 30-JUL-2002  
Homo sapiens 3 BAC RP11-170K4 (Roswell Park Cancer Institute Human  
BAC Library) complete sequence.



```

/rpt_family="AT_rich"
7595..7717
/rpt_family="MER5A"
complement(7788..8198)
/rpt_family="MLT11"
complement(8536..8591)
/rpt_family="L2"
9634..9936
/rpt_family="Alusx"
complement(10766..10814)
/rpt_family="MIR"
complement(10849..10964)
/rpt_family="MIR"
17399..17431
/rpt_family="(TTAAA)n"
complement(17432..17546)
/rpt_family="MIR"
17637..18073
/rpt_family="Charlielb"
18278..18313
/rpt_family="(TG)n"
21428..21841
/rpt_family="MSPD"
complement(21847..21926)
/rpt_family="LTR33A"
22387..22631
/rpt_family="MLT1G1"
complement(22768..23212)
/rpt_family="Tigger3(Golem)"
complement(23222..23279)
/rpt_family="Tigger3(Golem)"
23311..23437
/rpt_family="MLT1E2"
24105..24225
/rpt_family="MIR"
24226..24517
/rpt_family="MER33"
24518..24629
/rpt_family="MIR"
complement(24639..24724)
/rpt_family="L2"
25420..25994
/rpt_family="L2"
26024..26059
/rpt_family="(TG)n"
26565..27405
/rpt_family="L1MB8"
complement(27789..27855)
/rpt_family="L2"
28628..28778
/rpt_family="MER77"

Alignment Scores:
Pred. No.: 2,75e-163 Length: 167084
Score: 1936.00 Matches: 360
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-023-586B-2 (1-360) x AC092999 (1-167084)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyValLeu 20
|||||
Db 160921 ATGCTGTCATTTGCTTCTTCCAGGGGAGACAGAGCGGAGCCGCTGCTGGAGCTCTG 160980

Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGlu 40
|||||
Db 160981 CTCCTGGAGGGAGCCTCCCGGACATGGAGAAGCTGGACATGATACATCACAGAACAA 161040

Qy 41 GlyLeuGlnPheSerGluLysThrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
|||||
Db 161041 GGTCTGCCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 161100

Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80
|||||
Db 161101 ATCTTTATCTAGGCTGCCACATAAATGGCACCTGCTTGTGGCACCTCTCTGGGCCAAACC 161160

Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
|||||
Db 161161 AAGCGCTGGAGCTGTGCCACACCTATCTGTTGAACCTGATGTGTGGCCGACCTGCTTTAT 161220

Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpPheGlyGlu 120
|||||
Db 161221 GTGCTATTGCCCTTCTCATCATCATCTACTAGATCACAGTGGCCCTTCTGGGGAG 161280

Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
|||||
Db 161281 CTGCTGTCAAGCTGGTGACCTTCTCTTCTATATCAACCTTTACGGCAGCATCTGCTG 161340

Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
|||||
Db 161341 CTGACCTGCATCTCTGTGCACACCTGCTAGGTGTGTGCCACACCTGCTGTGCTGCC 161400

Qy 161 TyrArgThrArgArgHisAlaTyrLeuGlyThrSerThrThrTyrPheAlaLeuValValLeu 180
|||||
Db 161401 TACCGGACCGCAGGCATGCTGCTGGCGCACACACCTGCGGCGCTGGTGGTCTC 161460

Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200
|||||
Db 161461 CAGCTGCTGCCACACTGGCTGCTTCTCCACACGAGCTACATCAATGCCAGATGATCTGG 161520

Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
|||||
Db 161521 TATGACATGACAGCCCAAGAGAAATTTGTGCTGGCTTTTGCCTACGCCATAGTTCTGACA 161580

Qy 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
|||||
Db 161581 TTGCTGGCTTCTTCTCCCTCTTGGTCAATTTGGTGTGCTATTCATCTGATGGTCAGAG 161640

Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
|||||
Db 161641 CCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGCAACACACCGCCAGGTCAT 161700

Qy 261 ProAspHisProThrGlyValTyrProLeuHisProLeuPheCysAlaLeuProTyrHis 280
|||||
Db 161701 CCGGACCATCTACTGCTGTGTGGCTCTTTCACCTCTGTTTGTGGCTTCCATATCAC 161760

Qy 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
|||||
Db 161761 TCGCTCCTTCTACCTCACCATCTGCTTCTTCTTCTCAGACTGCCAGCTCTTGTATGGC 161820

Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
|||||
Db 161821 AGCCAGTGTGSCCTACAGATATGAGGCGCTCTGGTCACTGTGAGAGCTGCCCTCAACCC 161880

Qy 321 SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340
|||||
Db 161881 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGACTCAGCTCTCCAGAAACTGAGG 161940

Qy 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
|||||
Db 161941 CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAAGAGATGCCCGAGGTGTGAACAGATCT 162000

RESULT 6
AC021773 163958 bp DNA linear HTG 10-SEP-2000
LOCUS Homo sapiens clone RP11-12N17, WORKING DRAFT SEQUENCE, 15 unordered
DEFINITION pieces.
AC021773
VERSION AC021773.6 GI:10047750
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

1 (bases 1 to 163958)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens chromosome, clone RP11-12N17  
Unpublished  
2 (bases 1 to 163958)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,  
Choepeil,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,  
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,  
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,  
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Landers,T., Lehoczy,J., Levine,P., Liu,C., Liu,G., Locke,K.,  
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,  
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,  
Norman,C.H., O'Connor,T., O'Donnell,P., Oliviar,T.M., Peterson,K.,  
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,  
Roy,A., Santos,R., Severly,P., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,  
Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 10, 2000 this sequence version replaced gi:7248967.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L3302  
Center clone name: 12\_N17  
----- Summary Statistics  
Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 152766 bases at least Q40  
Consensus quality: 158075 bases at least Q30  
Consensus quality: 160495 bases at least Q20  
Insert size: 174000; agarose-fp  
Insert size: 162558; sum-of-contents  
Quality coverage: 4.5 in Q20 bases; agarose-fp  
Quality coverage: 4.9 in Q20 bases; sum-of-contents  
-----

TITLE  
JOURNAL  
COMMENT

\* 56887 65270: contig of 8384 bp in length  
\* 65271 65370: gap of 100 bp  
\* 65371 76394: contig of 11024 bp in length  
\* 76395 76494: gap of 100 bp  
\* 76495 100320: contig of 23826 bp in length  
\* 100321 100420: gap of 100 bp  
\* 100421 113127: contig of 12707 bp in length  
\* 113128 113227: gap of 100 bp  
\* 113228 126756: contig of 13529 bp in length  
\* 126757 126856: gap of 100 bp  
\* 126857 151061: contig of 24205 bp in length  
\* 151062 151161: gap of 100 bp  
\* 151162 163958: contig of 12797 bp in length.  
FEATURES  
Location/Qualifiers  
1. 163958  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-12N17"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. 7254  
/note="assembly\_fragment"  
clone\_end:sp6  
vector\_side:left  
7355..8612  
/note="assembly\_fragment"  
8713..11463  
/note="assembly\_fragment"  
11564..14507  
/note="assembly\_fragment"  
14608..22452  
/note="assembly\_fragment"  
22553..35086  
/note="assembly\_fragment"  
35187..44278  
/note="assembly\_fragment"  
44379..56786  
/note="assembly\_fragment"  
56887..65270  
/note="assembly\_fragment"  
65371..76394  
/note="assembly\_fragment"  
76495..100320  
/note="assembly\_fragment"  
100421..113127  
/note="assembly\_fragment"  
113228..126756  
/note="assembly\_fragment"  
126857..151061  
/note="assembly\_fragment"  
151162..163958  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right  
BASE COUNT 49230 a 32369 c 32422 g 48536 t 1401 others  
ORIGIN

Alignment Scores:  
Pred. NO.: 4.05e-162 Length: 163958  
Score: 1923.00 Matches: 360  
Percent Similarity: 99.72% Conservative: 0  
Best Local Similarity: 99.72% Mismatches: 0  
Query Match: 99.33% Indels: 1  
DB: 2 Gaps: 0  
US-10-023-586B-2 (1-360) x AC021773 (1-163958)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgGlySerArgGlyAlaLeu 20  
Db 98103 ATGCTGTCCATTTTCTCTCCAGGGGAGACAGAGCGGCGTGGAGCTCTG 98162  
Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
Db 98163 CTCCTGGAGGAGCCTCCCGGCACATGGAGAGGTGGACATCAATACATCAGGACAA 98222

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 15 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 7254: contig of 7254 bp in length  
\* 7255 7354: gap of 100 bp  
\* 7355 8612: contig of 1258 bp in length  
\* 8613 8712: gap of 100 bp  
\* 8713 11463: contig of 2751 bp in length  
\* 11464 11563: gap of 100 bp  
\* 11564 14507: contig of 2944 bp in length  
\* 14508 14607: gap of 100 bp  
\* 14608 22452: contig of 7845 bp in length  
\* 22453 22552: gap of 100 bp  
\* 22553 35086: contig of 12534 bp in length  
\* 35087 35186: gap of 100 bp  
\* 35187 44278: contig of 9092 bp in length  
\* 44279 44378: gap of 100 bp  
\* 44379 56786: contig of 12408 bp in length  
\* 56787 56886: gap of 100 bp

QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60  
 |||||  
 Db 98223 GGTCCTGCCAGTTCTCAGAGAAGTACAGCAAGTCTACTCTCCCTGGCCTACAGTATC 98282

QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80  
 |||||  
 Db 98283 ATCTTTATCTAGGGTGGCCTAAATGGCACTGTCTTGTGGCAGCTCTCTGGGGCCAAACC 98342

QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
 |||||  
 Db 98343 AAGCGTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCCGACCTGCTTTAT 98402

QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120  
 |||||  
 Db 98403 GTGCTATTGCCCTTCTCATCATCACTACTAGTACAGAGTGGCCCTTCGGGGAG 98462

QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
 |||||  
 Db 98463 CTGCTCTGAAGCTGGTGCACCTTCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCTG 98522

QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
 |||||  
 Db 98523 CTGACCTGCATCTCTGTGCACCAAGTTCTCAGGTGTGTGCCACCACTGTCTCGCTGCC 98582

QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180  
 |||||  
 Db 98583 TACCGGACCGCAGGATGCTGGTGGCCAGCAGCACCACTGGCCCTGGTGGTCTCTC 98642

QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200  
 |||||  
 Db 98643 CAGCTGTGCCACACAGTGGCTTCTCCACAGCACTATCAATGCCAGATGATCTGG 98702

QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220  
 |||||  
 Db 98703 TATGACATGACCAAGCAAGAAATTTTGTATGGCTTTTGTCCCTACCGCATAGTCTTGACA 98762

QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
 |||||  
 Db 98763 TTGCTCTGGCTTTCTTCCCTCCTTGTGTCTATTTTGGTGTGTCTATTCAGTGTCTAGGAG 98822

QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260  
 |||||  
 Db 98823 CCTGATCAAGCCAGAGAGAAACCTCATGAGGAGAGCAACACAGCCGAGCCAGGTCCAT 98882

QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280  
 |||||  
 Db 98883 CCGGACCATCTACTGGTGTGGCCCTCTTCACCCCTCTGTTTGTGGCCCTTCCATATCAC 98942

QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300  
 |||||  
 Db 98943 TCGCTCTCTTACCTCACCATCTGCTTCTCTGCTTTCTCAGGACTGCCAGCTCTTGATGCC 99002

QY 301 Ser-GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnGlnLeuProGlnPr 320  
 |||||  
 Db 99003 AGCCCAAGTGTGGCTCAAGATATGAGGGCTCTGGTGTAGTGTGAGCAGCTGCTCAACC 99062

QY 320 oSerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuLysLeuAr 340  
 |||||  
 Db 99063 CAGTCTGTACTTCTTTCAAGGGGGGCAAAATAGAGTCAGGCTCTCCAGAACTGAG 99122

QY 340 gGlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgse 360  
 |||||  
 Db 99123 GCAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCCGGGTTGACAGATC 99182

QY 360 r 360  
 |||||  
 Db 99183 T 99183

RESULT 7  
 AF411108  
 LOCUS  
 DEFINITION Homo sapiens GPR79 pseudogene, complete sequence. PRI 01-NOV-2001

AF411108  
 AF411108.1 GI:16566321  
 VERSION  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 1019)  
 AUTHORS Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhiko,O.,  
 Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.  
 TITLE Discovery and mapping of ten novel G protein-coupled receptor genes  
 JOURNAL Gene 275 (1), 83-91 (2001)  
 MEDLINE 21458557  
 PUBMED 11574155  
 REFERENCE  
 2 (bases 1 to 1019)  
 AUTHORS Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhiko,O.,  
 Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-AUG-2001) Department of Pharmacology, University of  
 Toronto, 8 Taddle Creek Rd. Rm. 4353, Toronto, Ont M5S 1A8, Canada  
 FEATURES  
 source  
 1..1019  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 gene  
 1..1019  
 /gene="GPR79"  
 /pseudo  
 BASE COUNT 200 a 308 c 252 g 259 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,06e-154 Length: 1019  
 Score: 1806.00 Matches: 339  
 Percent Similarity: 99.71% Conservative: 0  
 Best Local Similarity: 99.71% Mismatches: 0  
 Query Match: 93.29% Indels: 93 1  
 DB: 9 Gaps: 0  
 US-10-023-586B-2 (1-360) x AF411108 (1-1019)  
 QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20  
 |||||  
 Db 1 ATGCTCTCCATTGCTTCTTCCAGGGGAAGCAGAGCGGGAGCGTCTGGAGCTCTG 60  
 QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
 |||||  
 Db 61 CTCTCTGGAGGGACCTCCCGGGACATGGAGAGTGGACATGAATACATCACAGGACAA 120  
 QY 41 GlyLeuCysGlnPheSerGluTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60  
 |||||  
 Db 121 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180  
 QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80  
 |||||  
 Db 181 ATCTTTATCTAGGGTGGCCTAAATGGCACTGTCTTGTGGCAGCTCTCTGGGGCCAAACC 240  
 QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
 |||||  
 Db 241 AAGCGTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCCGACCTGCTTTAT 300  
 QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120  
 |||||  
 Db 301 GTGCTATTGCCCTTCTCATCATCACTACTAGTACAGAGTGGCCCTTCGGGGAG 360  
 QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
 |||||  
 Db 361 CTGCTCTGAAGCTGGTGCACCTTCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCTG 420  
 QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
 |||||  
 Db 421 CTGACCTGCATCTCTGTGCACCAAGTTCTCAGGTGTGTGCCACCACTGTCTCGCTGCC 480  
 QY 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180

```
Db 481 TACCGGACCCAGCAATGCTGGCTGGCCACAGCACCTGGGCTGGTGGCTCCT 540
Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
Db 541 CAGCTGCTGCCACACTGGCTTCTCCACAGGACTACATCAATGGCCAGATGATCGG 600
Qy 201 TyrAspMetThrSerGlnGlnAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 601 TATGACATGACCAAGCAAGAAATTTTCATCGGCTTTTGGCTACGGCATAGTCTGACA 660
Qy 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGln 240
Db 661 TTGCTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 720
Qy 241 ProAspGlnAlaArgGlyGlnProHisGlnAspArgGlnHisSerProSerGlnValHis 260
Db 721 CCTGATCATCAAGCAGGAGAACCTCATGAGGACAGGCAACACAGCCGAGCCAGTCCAT 780
Qy 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
Db 781 CCGGACCATCTACTGCTGCTGGCTTCTTCCACCTCTGTTTGTGGCTTCCATATCAC 840
Qy 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
Db 841 TCGCTCCTTCTACCTCAACATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360
Qy 301 Ser-GlnCysGlyLeuGlnAspMetGluAlaSerGlyGlnCysGlnGlnLeuProGlnPr 320
Db 901 AGCCAGTGTGGCTTACAAGATATGAGGAGCCCTCTGGTGGTGTGAGCAGCTGCCTCA 360
Qy 320 oSerProValLeuSerPheLeuGlyGlyLysAsnArgValArgLeuGlnLysLeu 339
Db 961 CAGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAGGCTCCTCCAGAACTG 1018

RESULT 8
AX458308 AX458308 1020 bp DNA linear PAT 08-JUL-2002
LOCUS
DEFINITION Sequence 3 from Patent EP1215214.
ACCESSION AX458308
VERSION AX458308.1 GI:21725002
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Fildock, M.D.
AUTHORS
TITLE Novel polypeptide
JOURNAL Patent: EP 1215214-A 3 19-JUN-2002;
Pfizer Limited (GB) ; PFIZER INC. (US)
FEATURES
Location/Qualifiers
source 1..1020
/db_xref="taxon:9606"
BASE COUNT 200 a 309 c 252 g 259 t
ORIGIN

Alignment Scores:
Pred. No.: 1,65e-152 Length: 1020
Score: 1785.00 Matches: 337
Percent Similarity: 99.12% Conservative: 1
Best Local Similarity: 98.83% Mismatches: 1
Query Match: 92.20% Indels: 2
DB: Gaps: 0

US-10-023-586B-2 (1-360) x AX458308 (1-1020)
Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
Db 1 ATGCTGTCCATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 60
Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGln 40
```

```
Db 61 CTCTCTGAGGAGAGCTCCCGGACATGGAGAGGTGGACATGAATATCATCAGAGAA 120
Qy 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
Db 121 GGTCTCTGCCAGTCTCTCAGAGAAGTACAGCAAGTCTACCTCTCTCCCTGGCCCTACAGTATC 180
Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTrpGlyGlnThr 80
Db 181 ATCTTTATCTCTAGGCTGCCACTAAATGGCACTGCTCTTGTGGCACTCTCTGGGGCAAAACC 240
Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
Db 241 AAGGCTGTGAGCTGTGCCACCATCTATCTGTGAACCTGATGGTGGCGGACCTGCTTTAT 300
Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTyrProPheGlyGlu 120
Db 301 GTGCTATTTGCCCTTCTCTCATCATCTACCTACTAGATGACAGGTGGCCCTTGGGGAG 360
Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
Db 361 CTGCTCTGCAAGCTGGTGCACCTTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 420
Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
Db 421 CTGACCTGCACTCTGTGGCACCATCTCTAGTGTGGTGGCACCCACTGTGTTCGTGGCC 480
Qy 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValLeu 180
Db 481 TACCGACCCGACGATGCTGTGGTGGCACACGACCATCTGGCCCTGGTGGCTCCT 540
Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
Db 541 CAGTGTGTGGCCACACTGGCTTCTCCACACGAGTACATCAATGGCCAGATGATCTGG 600
Qy 201 TyrAspMetThrSerGlnGlnAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
Db 601 TATGACATGACCCAGCAAGAAATTTGATCGCTTTTGGCTACGGCATAGTCTGACA 660
Qy 221 LeuSerGlyPheLeuSer--LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnG 240
Db 661 TTGCTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
Qy 240 LuProAspGlnAlaArgGlyGlnProHisGluAspArgGlnHisSerProSerGlnValH 260
Db 721 AGCTGTATCAAGCCAGAGAGAACCTCATGAGAGAGCAACACAGCCCGAGGATCC 780
Qy 260 isProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrH 280
Db 781 ATCCGGACCATCTCTACTGCTGTGGCTCTTTCACCTCTCTGTTTGTGGCCCTTCCATATC 840
Qy 280 isSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300
Db 841 ACTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 900
Qy 300 LysSerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP 320
Db 901 GCACCCAGTGTGGCTTACAGATATGAGGAGCTCTGTTGTGAGTGTGAGCAGCTGCTCAAC 960
Qy 320 roSerProValLeuSerPheLysGlyLysAsnArgValArgLeuGlnLysLeu 339
Db 961 CCAGTCTCTGATCTTCTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCTCCAGAACTG 1019

RESULT 9
AX375230 AX375230 1076 bp DNA linear PAT 01-MAR-2002
LOCUS
DEFINITION Sequence 20 from Patent WO0210387.
ACCESSION AX375230
VERSION AX375230.1 GI:19169948
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```





```
Db 241 ACCTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTC 300
Qy 129 LeuPheThrIleAsnLeuTyrclySerIleLeuLeuLeuThrCysIleSerValHisGln 148
Db 301 CTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGCTATCTCTGTGCACACG 360
Qy 149 PheLeuGlyValCysHisProLeuLeuCysSerLeuProTyrclyArgThrArgHisalaTrp 168
Db 361 TTCTCTAGGTGTGTGCCACCCACTGTGTTCGCTGCTACCGGACCCGCGAGCATGCTCTGG 420
Qy 169 LeuGlyThrSerThrThrTrpAlaLeuValLeuGlnLeuLeuLeuLeuProThrLeuAlaPhe 188
Db 421 CTGGGCACGACGACCATCTGGGCCCTGTGGTCTCCAGCTGCTGCCACACTGGCCCTTC 480
Qy 189 SerHisThrAspTyrclyIleAsnGlyGlnMetIleTrpTyrclyMetThrSerGlnGluAsn 208
Db 481 TCCACACGCACTACATCAATGTCAGATGATCTGATGATGATGATGATGATGATGATGATG 540
Qy 209 PheAspArgLeuPheAlaTyrclyIleValLeuThrLeuSerGlyPheLeuSerLeuLeu 228
Db 541 TTTGATCGGCTTTTGGCTACGGCAGTGTCTGACATTTCTGGCTTTCTTTCCCTCTCT 600
Qy 229 GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248
Db 601 GGTCAATTTGGTGTCTATCTACTGATGTCAGGAGCTGATCAGACGAGGAGAACT 660
Qy 249 HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp 268
Db 661 CATGAGGACAGCAACACAGCCGAGGATCCATCCGACCATCTACTGCTGTGTGG 720
Qy 269 ProLeuHisProLeuPheCysAlaLeuProTyrclyHisSerLeuLeuLeuProHisLeu 288
Db 721 CCTCTTCACTCTCTTTTGTGCTTCCATATCATCTGCTCTCTCTCTCTCTCTCTCTCT 780
Qy 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308
Db 781 CTTCTGCTTCTCAGGACTGCCACCTTGTGATGCCAGCCAGTGTGGCTTCAACAGATATG 840
Qy 309 GluAlaSer 311
Db 841 GAGGCTCT 849

RESULT 11
AC121579
LOCUS Mus musculus chromosome UNK clone RP23-255H8, WORKING DRAFT ;
DEFINITION Mus musculus chromosome UNK clone RP23-255H8, WORKING DRAFT ;
ACCESSION AC121579
VERSION AC121579.1 GI:20986651
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 209273)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL 1 (bases 1 to 209273)
REFERENCE 2 (bases 1 to 209273)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@genome.wustl.edu
----- Project Information -----
Center project name: M.BA025H08
----- Summary Statistics -----
Alignment Scores:
Pred. No.: 2,46e-92 Length: 209273
Score: 1154.50 Matches: 236
Percent Similarity: 74.37% Conservative: 31
Best Local Similarity: 65.74% Mismatches: 83
Query Match: 59.63% Indels: 10
DB: Gaps: 2

BASE COUNT 59147 a 42827 c 43921 g 62577 t 801 others
ORIGIN
```

```
Sequencing vector: M13; 0%
Chemistry: Dye-terminator; 100%
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 205908 bases at least Q40
Consensus quality: 206285 bases at least Q30
Consensus quality: 206424 bases at least Q20
Insert size: 177000; agarose-fp
Insert size: 208741; sum-of-contigs
Quality coverage: 15.45 in Q20 bases; agarose-fp
Quality coverage: 12.20 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

```
* 1 1957: contig of 1957 bp in length
* 1958 2057: gap of unknown length
* 2058 7480: contig of 5423 bp in length
* 7481 7580: gap of unknown length
* 7581 21224: contig of 13644 bp in length
* 21225 21324: gap of unknown length
* 21325 40793: contig of 19469 bp in length
* 40794 40893: gap of unknown length
* 40894 73232: contig of 32339 bp in length
* 73233 73332: gap of unknown length
* 73333 113562: contig of 40230 bp in length
* 113563 113662: gap of unknown length
* 113663 165371: contig of 51709 bp in length
* 165372 165471: gap of unknown length
* 165472 208277: contig of 42806 bp in length
* 208278 208377: gap of unknown length
* 208378 209273: contig of 896 bp in length.
```

```
FEATURES
        Location/Qualifiers
            1..209273
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /chromosome="UNK"
                /clone="RP23-255H8"
```

```
misc_feature
            1..1957
                /note="assembly_name:Contig18"
misc_feature
            2058..7480
                /note="assembly_name:Contig19"
misc_feature
            7581..21224
                /note="assembly_name:Contig20"
misc_feature
            21325..40793
                /note="assembly_name:Contig21"
misc_feature
            40894..73232
                /note="assembly_name:Contig22"
misc_feature
            73333..113562
                /note="assembly_name:Contig23"
misc_feature
            113663..165371
                /note="assembly_name:Contig24"
misc_feature
            165472..208277
                /note="assembly_name:Contig25"
misc_feature
            208378..209273
                /note="assembly_name:Contig16"
```

```
BASE COUNT 59147 a 42827 c 43921 g 62577 t 801 others
ORIGIN
```

```
Alignment Scores:
Pred. No.: 2,46e-92 Length: 209273
Score: 1154.50 Matches: 236
Percent Similarity: 74.37% Conservative: 31
Best Local Similarity: 65.74% Mismatches: 83
Query Match: 59.63% Indels: 10
DB: Gaps: 2
```

US-10-023-586b-2 (1-360) x AC121579 (1-209273)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20  
DB 37890 ATGCTGGCTTTATTTCTTCAGAGAACG-----AGAGTCGACGC 37934  
QY 21 LeuLeuGluGly-AlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGlu 40  
DB 37935 CATCTTCGAGGAGCCTCTTTAGGGATGGAGAGCTAGATATCATCTCTCAAGAGGTGC 37994  
QY 40 nGlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60  
DB 37995 TGGCGCTGCCACTCTCAGAGAGCTACACGAAGCTACCTATCATACCTACAGTCT 38054  
QY 60 eilePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrPheSerTrpGlyGln 80  
DB 38055 CATCTTATGCTAGGATGCCCTCAATGGGCTGTCTATGCTCTCTGGTGCACAAAC 38114  
QY 80 rLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeu 100  
DB 38115 CAAGCGCTGAGCTGTGCCACCATCTACCTGATGAACCTGGTGGCTGACCTGCTTTA 38174  
QY 100 rValLeu---LeuProPheLeuIleThrTyrSerLeuAspArgTrpProPheGln 119  
DB 38175 TATGCTAAGCTGCTCTGCTCATCATCATCTACCTACGGGTGGGTGACACCTGGCTCTTG 38234  
QY 119 yGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIle 139  
DB 38235 GGAGATGCTGTGAGGTGGTGGGTCTCTTCTACACCAACCTCTACAGCAGCATCT 38294  
QY 139 uLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSer 159  
DB 38295 GCTGCTGACCTGCTCTCTACACCGCTCTCTAGGAGTGGCAGCCTCTGAGCTCACT 38354  
QY 159 uProTyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuVal 179  
DB 38355 GTCTACCGAGCTACCGGCGCCCTGGCAGGAGCTGCTGCCACCTGGACCTGGTAGT 38414  
QY 179 lLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMet 199  
DB 38415 TCTCCAATTGCTGCCACTCTGCTCTCTCCACAGACTATGTAATGCCAGGTGAT 38474  
QY 199 eTrpTyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleVal 219  
DB 38475 CTGCTATGACACAGCCGCGGAGCTGTTGATCATTTCTGCGCTACAGCATAGTACT 38534  
QY 219 uThrLeuSerGlyPhe---LeuSerLeuGlyHisPheGlyValLeuPheThrAspGly 238  
DB 38535 GAGCTTATCTGGGTGTCTCTCCCTCTCTGATCATCTTGGGTGCTTCTCTGATGGT 38594  
QY 239 GlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGln 258  
DB 38595 TAGGAGCTCACAAATCCAGTGGAGGACCTCCCAACGACAGCCACACGCCGTGCCAA 38654  
QY 259 ValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuPro 278  
DB 38655 GTCTATTTCGAGCATTCTGCTGTGTGCTGCTCTTCCACTCTGTTTGTGCCCTTCCA 38714  
QY 279 TyrHisSerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeu 298  
DB 38715 CATTCGCCGATCTCTTACTGTATCTCCGCTCTCTGAACTCAGAGCTGCCAGCTCTT 38774  
QY 299 AspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGlnLeuPro 318  
DB 38775 GTGGCAGCCAGCATGCCCTACAGGCTCTAGGCTCTGTAGCATGAGCATGAGCTGCT 38834  
QY 319 GlnProSerProValLeuSerPheLysGlyLysAsnArgValArgLeuGlnLys 338  
DB 38835 CAATCCAGTCTGCTACTTC---TTGTACAGGGCAACACAGAGCCCACTCTCCAGAG 38892  
QY 339 LeuArgGlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysPro 355  
DB 38893 CTGAGACAAACAGGTGGGTGAGCACCCCTGCTGTGAAGGAGCACCT 38943

## RESULT 12

AC116265/c

## LOCUS

## DEFINITION

Rattus norvegicus clone CH230-18307, \*\*\* SEQUENCING IN PROGRESS

\*\*\*, 46 unordered pieces.

## ACCESSION

AC116265

## VERSION

HTG: HTGS\_PHASE1.

## KEYWORDS

Norway rat.

## SOURCE

Rattus norvegicus

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

## REFERENCE

1 (bases 1 to 91107)

## AUTHORS

Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,

Alsbrooks, S.L., Amaral, H.C., Are, J.R., Ayala, M., Banks, T.,

Barbata, J., Benton, J., Binge, K., Blankenburg, K., Bonnin, D.,

Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, N.P., Bryant, N.P.,

Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,

Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,

Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,

Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,

Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,

Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,

Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,

Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,

Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,

Gorell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,

Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,

Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,

Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,

Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,

Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,

Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,

Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Lousegh, H.,

Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,

Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,

Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,

Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,

Nguyen, N., Nickerson, E., Nwokoko, S., Ogih, M., Okuwono, G.,

Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,

Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,

Rivas, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Saverly, G.,

Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,

Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,

Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,

Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,

Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,

Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,

Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,

Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,

Weinstock, G., and Gibbs, R.

Direct Submission

Unpublished

2 (bases 1 to 91107)

Worley, K.C.

## TITLE

Direct Submission

Submitted (26-MAR-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 91107)

Worley, K.C.

## TITLE

Direct Submission

Submitted (18-JUL-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One

Baylor Plaza, Houston, TX 77030, USA

On Jul 14, 2002 this sequence version replaced gi:19718459.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GTRS  
 Center clone name: CH230-18307  
 ----- Summary Statistics  
 Sequencing vector: Plasmid;  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 36430 bases at least Q40  
 Consensus quality: 38793 bases at least Q30  
 Consensus quality: 40859 bases at least Q20  
 -----

\* NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently  
 consists of 46 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

```

1
* 1061: contig of 1061 bp in length
* 1161: gap of unknown length
* 1162: 2717: contig of 1556 bp in length
* 2718: 2817: gap of unknown length
* 2818: 4467: contig of 1650 bp in length
* 4468: 4567: gap of unknown length
* 4568: 6096: contig of 1529 bp in length
* 6097: 6196: gap of unknown length
* 6197: 7436: contig of 1240 bp in length
* 7437: 7536: gap of unknown length
* 7537: 8767: contig of 1231 bp in length
* 8768: 8867: gap of unknown length
* 8869: 10133: contig of 1266 bp in length
* 10134: 10233: gap of unknown length
* 10234: 11520: contig of 1287 bp in length
* 11521: 11620: gap of unknown length
* 11622: 13036: contig of 1416 bp in length
* 13037: 13136: gap of unknown length
* 13137: 14227: contig of 1091 bp in length
* 14228: 14327: gap of unknown length
* 14328: 15361: contig of 1034 bp in length
* 15362: 15461: gap of unknown length
* 15462: 16998: contig of 1537 bp in length
* 16999: 17098: gap of unknown length
* 17099: 18219: contig of 1121 bp in length
* 18220: 18319: gap of unknown length
* 18320: 19563: contig of 1244 bp in length
* 19564: 19663: gap of unknown length
* 19664: 21323: contig of 1660 bp in length
* 21324: 21423: gap of unknown length
* 21424: 22609: contig of 1186 bp in length
* 22610: 22709: gap of unknown length
* 22710: 24220: contig of 1511 bp in length
* 24221: 24320: gap of unknown length
* 24321: 25466: contig of 1146 bp in length
* 25467: 25566: gap of unknown length
* 25567: 27297: contig of 1731 bp in length
* 27298: 27397: gap of unknown length
* 27398: 28793: contig of 1396 bp in length
* 28794: 30289: contig of 1396 bp in length
* 30290: 31746: gap of unknown length
* 30390: 31846: gap of unknown length
* 31847: 33603: contig of 1757 bp in length
* 33604: 33703: gap of unknown length
* 33704: 35261: contig of 1558 bp in length
* 35262: 35361: gap of unknown length
* 35362: 37315: contig of 1954 bp in length
* 37316: 37415: gap of unknown length

```

```

* 37416 38787: contig of 1372 bp in length
* 38788 38887: gap of unknown length
* 38888 40559: contig of 1672 bp in length
* 40560 40659: gap of unknown length
* 40660 42655: contig of 1996 bp in length
* 42656 42755: gap of unknown length
* 42756 44162: contig of 1407 bp in length
* 44163 44262: gap of unknown length
* 44263 45600: contig of 2238 bp in length
* 45601 46500: gap of unknown length
* 46501 48572: contig of 1972 bp in length
* 48573 48672: gap of unknown length
* 48673 50544: contig of 1872 bp in length
* 50545 50644: gap of unknown length
* 50645 51883: contig of 1239 bp in length
* 51884 51983: gap of unknown length
* 51984 54568: contig of 2585 bp in length
* 54569 54668: gap of unknown length
* 54669 56245: contig of 1577 bp in length
* 56246 56345: gap of unknown length
* 56346 58742: contig of 2397 bp in length
* 58743 58842: gap of unknown length
* 58843 61978: contig of 3136 bp in length
* 61979 62078: gap of unknown length
* 62079 63933: contig of 1855 bp in length
* 63934 64033: gap of unknown length
* 64034 65836: contig of 1803 bp in length
* 65837 65936: gap of unknown length
* 65937 68259: contig of 2323 bp in length
* 68260 68359: gap of unknown length
* 68360 70942: contig of 2583 bp in length
* 70943 71042: gap of unknown length
* 71043 74577: contig of 3535 bp in length
* 74578 74677: gap of unknown length
* 74678 76364: contig of 1687 bp in length
* 76365 76464: gap of unknown length
* 76465 80009: contig of 3545 bp in length
* 80010 80109: gap of unknown length
* 80110 85962: contig of 5853 bp in length
* 85963 86062: gap of unknown length
* 86063 91107: contig of 5045 bp in length. ;

FEATURES             Location/Qualifiers
     source            1..91107
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-18307"

BASE COUNT   25705 a 18632 c 18796 g 23400 t 4574 others
ORIGIN
Alignment Scores:
Pred. NO.:      1.37e-90      Length:      91107
Score:          1130.00      Matches:      233
Percent Similarity: 74.72%      Conservative: 36
Best Local Similarity: 64.72%      Mismatches: 79
Query Match:     58.37%      Indels:      13
DB:              2              Gaps:      4

US-10-023-586B-2 (1-360) x AC116265 (1-91107)

Qy      1 MetLeuSerIleLeuLeuProSerArgGlySerArgGlySerArgGlyAlaLeu 20
      ||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 78493 ATGCTGTGCGTTCCTTATTCCTTACAGGAGCAGCAAGT-----AGAGGCCAT--- 78446

Qy      21 LeuLeu-GluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluG1 40
      ||| ||| ||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 78445 ---CTTTGAGGAAGCCTCCGTTGAGATGGAGAGCTACACGTCTCAGAGATGA 78389

Qy      40 nGlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerI1 60
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 78388 AGGCTTCTGCCACTTCTCGGAGAGCTACAGCAAGCTTACCTATCGCTAACCTACAGTCT 78329

Qy      60 e1lePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnTh 80
      :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

```

Db	78328	CATCTTCATGCTAGGAATGCCCTCAATAGGCGCGCTCTTATGGCTCTCTGGTGCCAAAC	78266
QY	80	rLysArgTrpSerCysAlaThrThrTrpLeuValAsnLeuMetValAlaAspLeuLeuTy	100
Db	78268	CAAGCCCTGGAGCTGTACCAACATCTACCTGATGAACCTGATGCTGGCGACCTGCTTTA	78209
QY	100	rValLeu---LeuProPheLeuLeuLeuThrTrpSerLeuAspAspArgTrpPropheGl	119
Db	78208	CATGCTCAGTCTGCCCTTCCTCATCATCACGCGCGGGTGACACCTGGCTCTTTGG	78149
QY	119	yGluLeuLeuCysLysLeuValHisPheLeuPheTyrlleAsnLeuTyrlGlySerlleLe	139
Db	78148	GGAGACGCTGTGCAGGCTGGTGGCTTTCCTGTTCTACCAACCTCTACAGCAGCATCT	78089
QY	139	uLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLe	159
Db	78088	GCTGCTCAGCTGCATCTCTATACCGCTTCTAGGAGTGTGCCACCCCTGAGCTCACT	78029
QY	159	uProTyArgThrArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValVa	179
Db	78028	GTCCCTTCGGACCCACCGCATCTGGCAAGCAGCGCCACCACCTGGACCCCTGGTAGT	77969
QY	179	lLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrlleAsnGlyGlnMetIl	199
Db	77968	CGTCCAATTGCTACCGACTCTGGTCTTCTCCACACAGACTATGTAATGCCACGTGAT	77909
QY	199	eTrpTyArgMetThrSerGlnGlnGluAsnPheAspArgLeuPheAlaTyrlGlylleValLe	219
Db	77908	CTGCTACGACACAGCCACCGAGAAATTCGATCATTTCTTCGCCCTACAGCATGGTGCT	77849
QY	219	uThrLeuSerGlyPhe--LeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGly	238
Db	77848	GACACTATCTGGGTTGTCTCCCTCCTTGATCATCTTCGGCTGCTACTCTCTGATGGT	77789
QY	239	GlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGln	258
Db	77788	CAGAGAGCTGACAATCCAGCAGAGGAGCTCCCAACACAGACGCCACGAGCCGCTGCCAA	77729
QY	259	ValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuPro	278
Db	77728	GTCTATCGGACCAATCTCTGGTGTGTAGCCCTCTTCACTGCTCTGTTTGTGCCCTTCCA	77669
QY	279	TyrHisSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeu	298
Db	77668	CATTGCCCGCTTCTTACTTTGGTCATCCGCTTCTCGAACCACACAGGATTGCCAGCTCTT	77609
QY	299	AspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGlyCysGluGlnLeuPro	318
Db	77608	GTTGGCAGCCAGCATGGCTCAAGATCTGTAGCCCTCTGTTAAGCGCGAGCAGCTGCC	77549
QY	319	GlnProSerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLys	338
Db	77548	CAACCCGGTCTGTACTCT--TTGTCAGGCGCAACACAGAGTCCCACTCTCTCCAGAG	77491
QY	339	LeuArgGlnAsnLysLeuGlyGluHisProAla-----GlyArgLysArg	353
Db	77490	CTGAGACAAAACAGGTGGTGAGCACCCCACTAGTGTGAAGGAGGACAGCTAGA	77437

RESULT	13
AX147768	
LOCUS	AX147768
DEFINITION	Sequence 13 from Patent WO0136473.
ACCESSION	AX147768
VERSION	AX147768.1 GI:14346805
KEYWORDS	.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 510) Vogeli,G., Wood,L.S., Parodi,L.A., Hiebsch,R.R., Lind,P., Slightom,J., Schellin,K.A., Kaytes,P.S., Bannigan,C.M., Ruff,V., Seitz,T. and Huff,R.M.

TITLE	Novel g protein-coupled receptors			
JOURNAL	Patent: WO 0136473-A 13 25-MAY-2001;			
FEATURES	PHARMACIA & UPJOHN COMPANY (US)			
source	Location/Qualifiers	1..510	/organism="Homo sapiens"	
BASE COUNT	92 a 156 c 126 g	136 t		
ORIGIN				
Alignment Scores:				
Pred. No.:	1..32e-74	Length:	510	
Score:	921.00	Matches:	170	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	47.5%	Indels:	0	
DB:	6	Gaps:	0	
US-10-023-586B-2 (1-360) x AX147768 (1-510)				
Qy	83	TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu	102	
Db	1	TGGAGCTGTGCCACCACTATCTGGTAACCTGATGTGGCCGACCTGCTTTATGTGCTA	60	
Qy	103	LeuProPheLeuIleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeu	122	
Db	61	TTGCCCTTCCATCATCACTACTACTAGATGACAGTGGCCCTTCGGGGAGCTGCTC	120	
Qy	123	CysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThr	142	
Db	121	TGCAAGCTGGTGCACCTTCCTGTTCTATATCAACCTTTACGGCAGCATCCTGCTGCTACC	180	
Qy	143	CysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg	162	
Db	181	TGCATCTCTGTGCACCACTTCTAGTGTGTGCACCACTGTGTGCTGCCCTACCGG	240	
Qy	163	ThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeu	182	
Db	241	ACCCGACGCATGGCTGGCTGGGACACCAACCACTGGGGCCCTGGTGTCTCCACGCTG	300	
Qy	183	LeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAsp	202	
Db	301	CTGCCACACTGGCTTCTCCACACGACTACATCAATGSCCAGATGATCTGGTATGAC	360	
Qy	203	MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSer	222	
Db	361	ATGACACCGCAAGAAATTTTGATCGGGCTTTTGGCTACGGCATAGTTCTGACATTGCT	420	
Qy	223	GlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAsp	242	
Db	421	GGCTTTCTTCCCTCTTGGTGCATTTTGGTGTGCTATTACTGATGTCAGGACCTTGAT	480	
Qy	243	GlnAlaArgGlyGluProHisGluAspArg	252	
Db	481	CAAGCCAGGAGACCTCATGAGCAGG	510	

RESULT 14					
AC120742/c					
LOCUS					
DEFINITION	AC120742	192524 bp	DNA	linear	HTG 24-JUL-2002
	Rattus norvegicus clone CH230-15014,	*** SEQUENCING IN PROGRESS			
	*** 56 unordered pieces.				
ACCENSION	AC120742				
VERSION	AC120742.3	GI:21908162			
KEYWORDS	HTG; HTGS_PHASE1.				
SOURCE	Rattus norvegicus.				
ORGANISM	Rattus norvegicus				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
	Rattus.				
REFERENCE	1 (bases 1 to 192524)				
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbaria,J., Benton,J., Bimace,K., Blankenburg,K., Bonnin,D.,				

Bouck, J., Bowie, S., Brieve, M., Brown, E., Brown, M., Bryant, N.P.,  
 Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
 Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
 Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
 Earnhart, C., Edgar, D., Edwards, C., Elhaj, C., Escotto, M.,  
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
 Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
 Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
 Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
 Homsi, F., Howard, S., Huber, J., Huiy, S., Hume, J., Jackson, L.E.,  
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
 Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
 Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H.,  
 Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
 Maheshwari, M., Mapua, P., Martin, K., Martindale, A., Martinez, E.,  
 Massey, E., Mahoney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
 Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,  
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, X.,  
 Rives, M., Rojas, A., Rojoubokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
 Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,  
 Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,  
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, K., Tamerisa, K., Tang, H.,  
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
 Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,  
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
 Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K.,  
 Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
 Weinstein, G. and Gibbs, R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 192524)  
 Worley, K.C.  
 Direct Submission  
 Submitted (09-MAY-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 192524)  
 Worley, K.C.  
 Direct Submission  
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 19, 2002 this sequence version replaced gi:20564367.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GXJ  
 Center clone name: CH230-15014  
 ----- Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 137507 bases at least Q40  
 Consensus quality: 143272 bases at least Q30  
 Consensus quality: 147627 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 56 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 1027: contig of 1027 bp in length  
 \* 1028 1127: gap of unknown length  
 \* 1128 1127: contig of 1024 bp in length  
 \* 2151: gap of unknown length  
 \* 2252 2251: gap of unknown length  
 \* 2252 2252: contig of 1271 bp in length  
 \* 3523 3523: gap of unknown length  
 \* 3623 3623: contig of 1443 bp in length  
 \* 5065 5065: gap of unknown length  
 \* 5165 5165: gap of unknown length  
 \* 5066 6860: contig of 1695 bp in length  
 \* 5166 6860: gap of unknown length  
 \* 6861 7995: contig of 1035 bp in length  
 \* 7996 8095: gap of unknown length  
 \* 8096 9381: contig of 1286 bp in length  
 \* 9382 9481: gap of unknown length  
 \* 9482 10544: contig of 1063 bp in length  
 \* 10545 10644: gap of unknown length  
 \* 10645 12080: contig of 1436 bp in length  
 \* 12081 12180: gap of unknown length  
 \* 12181 13574: contig of 1394 bp in length  
 \* 13575 13674: gap of unknown length  
 \* 13675 14766: contig of 1092 bp in length  
 \* 14767 14866: gap of unknown length  
 \* 14867 16102: contig of 1236 bp in length  
 \* 16103 16202: gap of unknown length  
 \* 16203 17487: contig of 1285 bp in length  
 \* 17488 17587: gap of unknown length  
 \* 17588 18875: contig of 1288 bp in length  
 \* 18876 18975: gap of unknown length  
 \* 18976 20257: contig of 1282 bp in length  
 \* 20258 21924: contig of 1567 bp in length  
 \* 21925 22024: gap of unknown length  
 \* 22025 23072: contig of 1048 bp in length  
 \* 23073 23172: gap of unknown length  
 \* 23173 24999: contig of 1827 bp in length  
 \* 25000 26426: gap of unknown length  
 \* 26427 26526: contig of 1327 bp in length  
 \* 26527 27831: contig of 1305 bp in length  
 \* 27832 27931: gap of unknown length  
 \* 27932 29816: contig of 1885 bp in length  
 \* 29817 29916: gap of unknown length  
 \* 29917 31690: contig of 1774 bp in length  
 \* 31691 31791: gap of unknown length  
 \* 31791 34578: contig of 2788 bp in length  
 \* 34579 34678: gap of unknown length  
 \* 34679 36533: contig of 1855 bp in length  
 \* 36534 36633: gap of unknown length  
 \* 36634 38324: contig of 1691 bp in length  
 \* 38325 38424: gap of unknown length  
 \* 38425 40498: contig of 2074 bp in length  
 \* 40499 40598: gap of unknown length  
 \* 40599 43132: contig of 2534 bp in length  
 \* 43133 43232: gap of unknown length  
 \* 43233 46098: contig of 2866 bp in length  
 \* 46099 46198: gap of unknown length  
 \* 46199 48303: contig of 2105 bp in length  
 \* 48304 48403: gap of unknown length  
 \* 48404 51105: contig of 2703 bp in length  
 \* 51107 51206: gap of unknown length  
 \* 51207 54222: contig of 3016 bp in length  
 \* 54223 54323: gap of unknown length  
 \* 54323 56846: contig of 2524 bp in length  
 \* 56847 56946: gap of unknown length  
 \* 56947 60999: contig of 4053 bp in length  
 \* 61000 61099: gap of unknown length  
 \* 61099 65048: contig of 3949 bp in length  
 \* 65049 65148: gap of unknown length

65149 67208: contig of 2060 bp in length  
67209 67308: gap of unknown length  
70014: contig of 2706 bp in length  
70114: gap of unknown length  
75273: contig of 5159 bp in length  
75274 75373: gap of unknown length  
75374 78558: contig of 3185 bp in length  
78559 78659: gap of unknown length  
81049: contig of 2391 bp in length  
81149: gap of unknown length  
84528: contig of 3779 bp in length  
84929 85028: gap of unknown length  
85029 89226: contig of 4198 bp in length  
89227 94002: gap of unknown length  
94003 94102: gap of unknown length  
94103 98447: contig of 4345 bp in length  
98448 98547: gap of unknown length  
98548 102392: contig of 3745 bp in length  
102393 106997: gap of unknown length  
106998 107097: gap of unknown length  
111457: contig of 4360 bp in length  
111458 111557: gap of unknown length  
111558 115579: contig of 4022 bp in length  
115580 119362: contig of 4283 bp in length  
119363 120062: gap of unknown length  
120063 126366: contig of 6304 bp in length  
126367 126466: gap of unknown length  
126467 133796: contig of 7330 bp in length  
133797 133896: gap of unknown length  
133897 142152: contig of 8256 bp in length  
142153 142252: gap of unknown length  
142253 151202: contig of 8950 bp in length  
151203 151302: gap of unknown length  
151303 158779: contig of 7477 bp in length  
158780 158879: gap of unknown length  
158880 168094: contig of 9215 bp in length

Alignment Scores:  
Pred. No.: 3,128-68 Length: 192524  
Score: 888.00 Matches: 176  
Percent Similarity: 75.5% Conservative: 28  
Best Local Similarity: 65.1% Mismatches: 57  
Query Match: 45.8% Indels: 9  
DB: 2 Gaps: 3

US-10-023-586B-2 (1-360) x AC120742 (1-192524)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerArgArgGlyAlaLeu 20  
Db 159732 ATGCTGGGTTTATTTCTTACAGGAGCAGCAGAGT-----AGAGCCAT--- 159685

QY 21 LeuLeu-GluGlyAlaSerArgMetGluLysValAspMetAsnThrSerGlnGlu 40  
Db 159684 ---CTTTGAGGAGCCCTCGGTTGAGATGGAGAGCTAGACGTGAATGCTCAGAGATGA 159628

QY 40 nGlyLeuCysGlnPheSerGlnLysThrLysGlnValTyrLeuSerLeuAlaTyrSer 60  
Db 159627 AGGCTTCGCCATCTTCGGAGAGCTACAGCAAGTCTACCTACCTACCTACCTACCT 159568

QY 60 eilePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrPheSerTyrGlnThr 80  
Db 159567 CATCTTCATGCTAGGATGCGCCCTCAATGGCGCGTCTTATGGCTCTCTGGTGCACAA 159508

QY 80 rLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeu 100  
Db 159507 CAAGCCCTGGAGCTGTACCACTACCTACCTACCTACCTACCTACCTACCTACCTACCT 159448

QY 100 rValLeu---LeuProPheLeuIleThrTyrSerLeuAspArgTrpProPheG 119  
Db 159447 CATGCTGAGTCTGCCCTTCTCATCATCATCATCATCATCATCATCATCATCATCATCAT 159388

QY 119 yGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLe 139  
Db 159387 GGAGAGCTGTGCAGCTGGTTCCTGTCTACACCACTCTACAGCAGCATCCT 159328

QY 139 uLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLe 159  
Db 159327 GCTGCTGAGCTGCATCTCTATACCGCTTTCTAGAGTGTGCCACCCCTGAGCTCACT 159268

QY 159 uProTyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValVa 179  
Db 159267 GTCCCTTCGGGACCCACCGCATACCTGGCAAGCAGCGCCACCATGGACCTGGTAGT 159208

QY 179 lLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIl 199  
Db 159207 CGTCCAATTGCTACCGACTCTGGTCTTCTCCACACAGACTATGTAATGGCCACGTGAT 159148

QY 199 eTrpTyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLe 219  
Db 159147 CTGCTACGACACAGCCAGCCAGAGAAATTTCTGATCATTTCTCGCCTACAGCATGGTCT 159088

QY 219 uThrLeuSerGlyPhe--LeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGly 238  
Db 159087 GACACTATCTGGGTTGTCTTCCCTCTTGATCATTTCTCGGCTGCTACTCTCTGATGGT 159028

QY 239 GlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGln 258  
Db 159027 CAGGACCTTGACAAATCCAGCAGGACCTCCCAAGCAGACAGCCGCGTGCCTGCA 158968

QY 259 ValHisProAspHisProThrGly 266  
Db 158967 GTCTATCCGGACCATCTGCTGGT 158944

RESULT 15  
AX244718/c  
LOCUS AX244718 681 bp DNA linear PAT 28-SEP-2001  
DEFINITION Sequence 47 from Patent WO0166750.  
ACCESSION AX244718  
VERSION AX244718.1 GI:15859597  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 681)  
AUTHORS Vogeli, G. and Wood, L.S.  
TITLE G protein-coupled receptors  
JOURNAL Patent: WO 0166750-A 47 13-SEP-2001;  
PHARMACIA & UPJOHN COMPANY (US)  
FEATURES  
Location/Qualifiers  
source 1..681  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 163 a 171 c 171 g 176 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2,596-57 Length: 681  
Score: 732.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 37.81% Indels: 0  
DB: 6 Gaps: 0

US-10-023-586B-2 (1-360) x AX244718 (1-681)

QY 229 GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248  
Db 681 GGTCACTTTGGTGTGCTATCTACTGATGTCAGAGCCTGATCAAGCCAGAGAACCT 622

QY 249 HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp 268  
Db 621 CATGAGGACAGGCAACACAGCCGAGCGAGGTCCATCCGACCATCTCTACTGCTGTGG 562

Qy 269 ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeu 288  
|||||  
Db 561 CCTCTTCAACCTCTGTTTGTGGCCCTTCCATATCACTCGCTCCTTCTACCTCACCATCTG 502  
  
Qy 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308  
|||||  
Db 501 CTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCCTACAAGATATG 442  
  
Qy 309 GluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPheLysGly 328  
|||||  
Db 441 GAGGCCCTGTGTGAGTGTGAGCAGCTGCTCAACCCAGTCTGTACTTTCTTCAAGGGG 382  
  
Qy 329 GlyLysAsnArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisPro 348  
|||||  
Db 381 GGCAAAATAGAGTCAGGCTCTCTCCAGAAACTGAGGCAGACAAAGTTGGGTGAGCATCCA 322  
  
Qy 349 AlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
|||||  
Db 321 GCTGGGAGGAAGAGATGCCAGGGTTGAACAGATCT 286

Search completed: May 27, 2003, 00:08:39  
Job time : 3925 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 21:29:43 ; Search time 304 Seconds  
(without alignments)  
2666.840 Million cell updates/sec

Title: US-10-023-586B-2

Perfect score: 1936

Sequence: 1 MSLILPSRSGSRGAL.....QNKLGHPAGRRCPCLNRS 360

oring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -dbv=xlh  
-O=/cgn2\_1/USPTO.spool/US10023586/runat\_22052003\_114849\_28633/app\_query.fasta.1.519  
-DB=N\_Geneseq\_101002 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10023586.ecgn.1.1.263.0runat\_22052003\_114849\_28633 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOB=6 -FGAPEXT=7  
-YGAPOB=10 -YGAPEXT=0.5 -DELOB=6 -DELEXT=7

Database :

N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1936	100.0	3143	24	AAS17746	Human genomic clon
2	1898	98.0	1063	24	AAS98135	Human DNA for pte
3	1798	92.9	1002	22	AH73516	Human G-protein-co
4	1785	92.2	1076	24	AAD29667	Human G-protein co
5	1541	79.6	850	24	AAS17747	Human P2Y-like G p
6	921	47.6	510	22	AAS50975	Human ngPCR12 codi
7	732	37.8	681	22	AAS30774	Human cDNA encodin
8	681	35.2	585	22	AAS30783	Human cDNA encodin
9	453	23.4	1163	22	AAD04981	Human purinergic r
10	445.5	23.0	1429	18	AAT74321	Human P2Y4 recepto
11	442	22.8	2025	22	AAD04980	Human purinergic r
12	420.5	21.7	984	17	AT18367	Mouse pancreas G-p
13	416	21.5	1842	16	AAO88134	Human P20 receptor
14	415	21.4	2138	24	AAD32937	Mus musculus GPCR
15	415	21.4	2138	24	AAD32944	Mus musculus GPCR
16	413	21.3	984	17	AT18368	Human placenta G-p
17	402	20.8	984	18	AAT70000	P20 purinergic rec
18	402	20.8	984	18	AAT747730	Human placenta pur
19	370	19.1	551	22	AAS42837	Human G Protein-Co
20	328	16.9	6721	24	AAS18599	Purinergic recepto
21	327.5	16.9	6721	24	AAS18600	Purinergic recepto
22	294	15.2	450	21	AA42673	Human secreted exp
23	292	15.1	234	22	AAS30782	Human cDNA encodin
24	281.5	14.5	1020	22	AAS51011	Human ngPCR54 codi
25	278.5	14.4	1011	24	AAL43942	Human G-protein-co
26	278.5	14.4	1014	22	AAS07948	Human cDNA encodin
27	278.5	14.4	1014	24	ABO78847	Human G-protein co
28	278.5	14.4	1014	24	ABN85630	Human P2Y-like rec
29	278.5	14.4	1014	24	AAD34278	Human AXOR89 (G-pr
30	278.5	14.4	1014	24	ABK11381	Human DNA encoding
31	278.5	14.4	1014	24	AAG98323	Human purinergic-r
32	278.5	14.4	1014	24	AAD26370	Human G-protein co
33	278.5	14.4	1288	24	ABL56197	Human P2Y1-li enco
34	278.5	14.4	1729	22	AAS08362	Human cDNA encodin
35	278.5	14.4	1729	23	ABV24026	Human prostate exp
36	278.5	14.4	1729	23	ABV25767	Human prostate exp
37	278.5	14.4	1729	23	ABV29909	Human prostate exp
38	278.5	14.4	1729	23	ABV30024	Human prostate exp
39	278.5	14.4	5435	24	AAK56198	Human P2Y1-li enco
40	278.5	14.4	9905	24	AAK98324	Human purinergic-r
41	278	14.4	1313	22	AAK52430	Human polynucleoti
42	275	14.2	1301	17	AAT33905	Human G-protein co
43	275	14.2	1301	21	AA70771	Human G-protein co
44	275	14.2	1301	21	AA260119	Human G-protein co
45	269	13.9	2162	19	AAV61993	Human mu-opioid re

ALIGNMENTS

RESULT 1

AAS17746

ID AAS17746 standard; DNA; 3143 BP.

XX AAS17746;

AC AAS17746;

XX 26-FEB-2002 (first entry)

DT Human genomic clone for P2Y-like G protein-coupled receptor.

DE Human; ds; P2Y-like G protein-coupled receptor; GPCR; COPD;

XX Chronic obstructive pulmonary disease; nervous system disease;

KW Parkinson's disease; multiple sclerosis; dementia; stroke;

KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;

KW bacterial infection; fungal infection; protozoan infection;

KW viral infection; pain; cancer; anorexia; bulimia; asthma;

KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;

KW angina pectoris; myocardial infarction; ulcer; inflammation; allergy;

psychotic disorder; neurological disorder; anxiety; schizophrenia;  
manic depression; delirium; severe mental retardation; dyskinesia.

Homo sapiens.

Key Location/Qualifiers  
CDS 520..2498  
/\*tag= a

/product= "P2Y-like GPCR"

W0200185764-A2.

15-NOV-2001.

09-MAY-2001: 2001WO-EP05244.

11-MAY-2000: 2000US-203582P.

21-FEB-2001: 2001US-269857P.

(FARB ) BAYER AG.

Ramakrishnan S;

WPI: 2002-075242/10.

P-PSDB; AA011251.

New polynucleotides for producing P2Y-like G protein-coupled receptors (GPCR) that are used for screening inhibitors or regulators of human P2Y-like GPCR, especially useful for treating pain, cancer or neurological disorders

Disclosure: Fig 1: 114pp; English.

The invention relates to an isolated polynucleotide encoding a P2Y-like G protein-coupled receptor (GPCR) polypeptide, its fragment, derivative or allele, a host cell containing an expression vector comprising the polynucleotide and screening for agents that regulate the GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR polypeptide, which may be employed for screening agents that inhibit or regulate human P2Y-like GPCR. The reagent or inhibitor of the human P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR disorders, particularly COPD (chronic obstructive pulmonary disease), peripheral or central nervous system disease (e.g. Parkinson's disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and many other diseases and disorders listed in the specification), benign prostatic hyperplasia or urinary incontinence. A pharmaceutical composition containing the modulators and/or regulators of P2Y-like GPCR is useful for modulating the activity of a P2Y-like GPCR.

In particular, these are useful for treating, preventing or ameliorating infections (e.g. bacterial, fungal, protozoan or viral infections), pain, cancer, anorexia, bulimia, asthma, acute heart failure, hypotension, hypertension, osteoporosis, diabetes, angina pectoris, myocardial infarction, ulcers, inflammation, allergies, psychotic or neurological disorders (e.g. anxiety, schizophrenia, manic depression, delirium, severe mental retardation or dyskinesias). The present sequence is a genomic clone encoding the P2Y-like GPCR of the invention.

Sequence 3143 BP; 818 A; 771 C; 726 G; 828 T; 0 other;

#### Alignment Scores:

Pred. No.:	1-29e-174	Length:	3143
Score:	1936.00	Matches:	360
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-023-586B-2 (1-360) x AAS17746 (1-3143)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
Db ATGCTGTCCATTTTCTTCCAGGGGAGCAGCAGGCGGCGTCTGTGGAGCTCTG 495

Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
Db CTCCTGGAGGAGCCCTCCGGGACATGGAGAAGTGGACATGAATACATCAGAGAACAA 555  
Qy 41 GlyLeuCysGlnPheSerClnLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60  
Db GGTCTCTGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGCGCTACAGTATC 615  
Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThr 80  
Db ATCTTTATCTTAGGCTGCCACTAAATGGCACTGCTGTGGCACTCTCTGGGGCCAAACC 675  
Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
Db AAGCGCTGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGTGGCCGACCTGTTAT 735  
Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrpProPheGlyGlu 120  
Db GTGCTATTGCGCTTCTCATCATCACTACTACTACATGATGACAGGTGGCCCTTCGGGGAG 795  
Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
Db CTGCTCTGCAAGCTGTGTCACCTCTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 855  
Qy 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
Db CTGACCTGCATCTCTGTGCACCACTTCTGTAGGTGTGTGCCACCACTGTGTTCGCTGCC 915  
Qy 161 TyrArgThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeu 180  
Db TACCGGACCGGAGGAGCTGCTGGCTGGCCACCACTGCGCTGGCCCTGCTGTGCTCTC 975  
Qy 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200  
Db CAGCTGTCTGCCACACTGGCTTCTCCACACGGACTACATCAATGGCCAGATGATCTGG 1035  
Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220  
Db TATGACATGACCAAGAGAGAAATTTTGTATCGCTTTTGTGCTAGCGCATAGTCTTGACA 1095  
Qy 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
Db TTGCTCTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1155  
Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260  
Db CCTGATCAAGCCAGAGAGAGAACTCATGAGGACAGCAACACACAGCCAGCGAGTCCAT 1215  
Qy 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280  
Db CCGGACCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1275  
Qy 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300  
Db TCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1335  
Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320  
Db AGCCAGTGTGGCTTACAAGATATGAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1395  
Qy 321 SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340  
Db AGTCTCTGTACTTCTTCTTCAAGGGGGGCAAAATAGATGAGTCAAGCTCTCTCAGAACTGAGG 1455  
Qy 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
Db CAGAACAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCCAAGGTTGACAGATCT 1515  
RESULT 2  
AAS98135  
ID AAS98135 standard; DNA; 1063 BP.  
XX  
AC AAS98135;

XX 12-MAR-2002 (first entry)  
 XX Human DNA for potential G protein-coupled receptor #92.  
 XX  
 XX Human; G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;  
 XX Alzheimer's disease; amyotrophic lateral sclerosis; asthma;  
 XX atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;  
 XX chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;  
 XX depression; epilepsy; macular degeneration; lymphoma; melanoma;  
 XX multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;  
 XX psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;  
 XX tuberculosis; cognition disorder; memory disorder; anorexia;  
 XX hormonal release disorder; cardiovascular activity disorder;  
 XX pain perception disorder; obesity; diabetes; obesity;  
 XX diabetes; hyperlipidaemia; stroke; gene therapy.  
 XX Homo sapiens.  
 XX WO200185791-A1.  
 XX  
 XX 15-NOV-2001.  
 XX  
 XX 11-MAY-2001; 2001WO-US15332.  
 XX  
 XX 11-MAY-2000; 2000US-203217P.  
 XX 18-MAY-2000; 2000US-205945P.  
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.  
 XX  
 XX Brown JP, Miller M, Burner G, Fabre-Suver C, Pritchard D;  
 XX WPI; 2002-066595/09.  
 XX  
 XX Novel G protein-coupled receptor polypeptides including galanin  
 XX receptor polypeptides useful for identifying modulators that are useful  
 XX for treating Alzheimer's disease, psoriasis, melanoma, multiple  
 XX sclerosis, stroke  
 XX  
 XX Claim 2; Page 98; 144pp; English.  
 XX  
 XX The invention relates to an isolated polypeptide encoded by a  
 XX nucleic acid molecule that is at least 80% identical to the G  
 XX protein-coupled (GPCR) polynucleotides included in the specification.  
 XX Also included are probes based on the GPCR sequences (including  
 XX antisense probes), a host cell comprising an expression vector comprising  
 XX the GPCR sequence, antibodies raised against the polypeptides,  
 XX and methods of identifying modulators of the polypeptides. The  
 XX polypeptides are useful for identifying modulator compounds which  
 XX function as modulators, activators, repressors, agonists or antagonists  
 XX of the novel GPCR polypeptides including the GAL4 polypeptide. The  
 XX antibodies and nucleic acid probes as described above can be used to  
 XX detect the presence of the polypeptides and nucleic acids and are used to  
 XX diagnose a variety of diseases or disorders in which GPCRs are involved  
 XX e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma,  
 XX atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy,  
 XX chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease,  
 XX depression, epilepsy, macular degeneration, lymphoma, melanoma,  
 XX multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease,  
 XX psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis,  
 XX tuberculosis and many other diseases listed in the specification. The  
 XX probes and antibodies are also useful for diagnosing cognition and memory  
 XX disorders, anorexia, hormonal release disorders, cardiovascular activity  
 XX disorders, pain perception disorders, obesity, diabetes, Alzheimer's  
 XX disease. Preferably, compounds that decrease or increase  
 XX the expression of galanin receptor (GAL4) can be used to treat obesity,  
 XX diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is  
 XX useful for treating the above mentioned disorders by gene therapy  
 XX techniques. The present sequence is a novel GPCR polynucleotide of the  
 XX invention.

XX Sequence 1063 BP; 216 A; 315 C; 272 G; 260 T; 0 other;  
 XX

Alignment Scores: 1.17e-171 Length: 1063  
 Pred. No.: 1898.00 Matches: 354  
 Score: 99.72% Conservative: 0  
 Percent Similarity: 99.72% Mismatches: 0  
 Best Local Similarity: 99.72% Indels: 1  
 Query Match: 98.04% Gaps: 0  
 DB: 24  
 US-10-023-586B-2 (1-360) x AAS98135 (1-1063)  
 QY 7 ProSerArgGlySerArgSerGlySerArgGlyAlaLeuLeuGluGlyAlaSer 26  
 DB 1 CCTTCCAGGGAAGCAGAGCGGAGCGCTGCTGCTCTCTGAGAGGAGGCTCC 60  
 QY 27 ArgAspMetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeuCysGlnPheSer 46  
 DB 61 CGGACATGGAGAGGTGGACATGATACATACAGGAACAAGGTCTCTGCCAGTTCTCA 120  
 QY 47 GluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerLeuPheLeuLeuGlyLeu 66  
 DB 121 GAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATCATCTTTATCTAGGGCTG 180  
 QY 67 ProLeuAsnGlyThrValLeuTyrHisSerTrpGlyGlnThrLysArgTrpSerCysAla 86  
 DB 181 CCACATAATGGCACGTGCTTTGTGGCACTCTCGGGGCCAAACCAAGCGCTGAGCTGCC 240  
 QY 87 ThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeuLeuProPheLeu 106  
 DB 241 ACCACCTATCTGGTGAACCTGATGGTGGCGGACCTGCTTTATGCTGCTATTTGCTCTC 300  
 QY 107 IleIleThrTyrSerLeuAspAspArgTrpProPheGlyGluLeuLeuCysLysLeuVal 126  
 DB 301 ATCATCACCTACTCATCTAGATGACAGGTGCCCTTCGGGAGCTGCTCTGCAAGCTGGT 360  
 QY 127 HisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerVal 146  
 DB 361 CACTTCTCTGCTTATATCAACCTTTACGGCAGCATCTCTGCTGACCTGCATCTCTGTG 420  
 QY 147 HisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArgTrpArgHis 166  
 DB 421 CACCAGTTCCTAGGTGTGGCCACCCACTGTGTGCTGCTACCGGACCCGCGAGCAT 480  
 QY 167 AlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeu 186  
 DB 481 GCCTGGTGGCCACAGCACCACTGGTGGCCCTGGTGGTCTCTCCAGCTGCTGCCACACTG 540  
 QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206  
 DB 541 GCCTTCTCCACACGAGTACTCATATGGCCAGATGATCTGGTATGACATGACCCAGCAA 600  
 QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSer 226  
 DB 601 GAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCTGACATGTTCTGGCTTTCTTTCC 660  
 QY 227 LeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGly 246  
 DB 661 CTCCTTGGTCATTTTGGTGTGCTATTCTACTGATGGTGGAGCGCTGATCAAGCCAGAGA 720  
 QY 247 GluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGly 266  
 DB 721 GAACCTCATGAGGACAGGCAACACAGCCCGAGCGGTCTCATCCGGACCATCTCTACCTG 780  
 QY 267 ValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHis 286  
 DB 781 GTGGGCTCTTCAACCTCTGTTTGTGGCCTTCCATATCATCTGCTCTCTTCTACCTCAC 840  
 QY 287 HisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer-GlnCysGlyLeuGln 306  
 DB 841 CATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGGCAGCCAGCTGTGGCCCTACA 900  
 QY 306 nAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPhe 326  
 DB 901 AGATATGGAGGCGCTCTGGTGAGTGTGAGAGCTGCCTCAACCCAGCTCTGTACTTTCTTT 960

QY 326 etysGlyGlyLysAsnArgValArgLeuGlnLysLeuArgGlnAsnLysLeuGlyG1 346  
 Db 961 CAAGGGGGCAAAATAGAGTCAAGCTCTCCAGAAACTGAGGCAGACAAAGTTGGGTGA 1020  
 QY 346 uHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
 Db 1021 GCATCCAGCTGGAGGAGAGAGATGCCAGGGTTGAACAGATCT 1063

RESULT 3  
 AAH73516

ID AAH73516 standard; DNA; 1002 BP.

AC AAH73516;

DT 25-SEP-2001 (first entry)

DE Human G protein-coupled receptor GPRV71 coding sequence.

XX  
 KW Human; guanosine triphosphate binding protein-coupled receptor;  
 KW G protein-coupled receptor; GPRv8; GPRV12; GPRV16; GPRV21; GPRV40;  
 KW GPRV47; GPRV51; GPRV71; GPRV72; cancer; liver cirrhosis;  
 KW Alzheimer's disease; cytostatic; hepatotropic; neurotropic;  
 KW neuroprotective; gene therapy; peptide therapy; ds.

XX Homo sapiens.

PN WO200148188-A1.

PD 05-JUL-2001.

XX 28-DEC-2000; 2000WO-JP09408.

XX 28-DEC-1999; 99JP-0375152.

PR 31-MAR-2000; 2000JP-0101339.

XX (HELI-) HELIX RES INST.

PI Matsumoto.S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;

PI Sugiyama T, Kishimoto T, Kanzaki K, Yasuda S, Inoue Y;

XX WPI: 2001-425662/45.

DR P-PSDB; AAG64125.

XX New DNA encoding guanosine triphosphate binding protein coupled  
 PT receptors and their expression products for screening potential  
 PT anticancer and neurotropic drugs and in diagnosis of these diseases

XX Example 1; Page 142-143; 170pp; Japanese.

XX The invention relates to nine human guanosine triphosphate binding  
 CC protein (G protein)-coupled receptors designated GPRv8, GPRV12, GPRV16,  
 GPRV21, GPRV40, GPRV47, GPRV51, GPRV71 and GPRV72, and to the  
 CC genes encoding them. These genes and proteins and antibodies against  
 the protein are useful in the treatment, prevention, diagnosis and  
 CC investigation of diseases associated with G protein-coupled receptors,  
 CC including cancer, cirrhosis of the liver and Alzheimer's disease.  
 CC The present sequence is a G protein-coupled receptor coding sequence  
 CC of the invention.

XX SQ Sequence 1002 BP; 207 A; 293 C; 249 G; 253 T; 0 other;

Alignment Scores:

Pred. No.: 3.85e-162 Length: 1002  
 Score: 1798.00 Matches: 331  
 Percent Similarity: 99.70% Conservative: 0  
 Best Local Similarity: 99.70% Mismatches: 1  
 Query Match: 92.87% Indels: 0  
 DB: 22 Gaps: 0

US-10-023-586B-2 (1-360) x AAH73516 (1-1002)

QY 29 MetGluLysValAspMetAsnThrSerGlnGluGlyLeuCysGlnPheSerGluLys 48

Db 1 ATGGAGAAGGTGGACATGAATACATCACAGGAACAAGGTCTCTGCAGGTTCTCAGAGAAG 60  
 QY TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu 68  
 Db 61 TACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATCATCTTTATCTAGGGGTGCACATA 120  
 QY AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrTrp 88  
 Db 121 AATGCCACTGTCTTGTGCACCTTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACCACC 180  
 QY TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeuLeuProPheLeuIleIle 108  
 Db 181 TATCTGGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCTATTGCTCCCTTCCTCATCATC 240  
 QY ThrTyrSerLeuAspAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHisPhe 128  
 Db 241 ACCTACTCAGTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCATTC 300  
 QY LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148  
 Db 301 CTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTGCTGACCTGCATCTCTGTGCACCAG 360  
 QY PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgArgHisAlaTrp 168  
 Db 361 TTCTTAGGTGTGCCACCCACTGTGTCTGCTGCCCTACCGGACCCGAGGATGCCTGG 420  
 QY LeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe 188  
 Db 421 CTGGCACCAGCACCACCTGGCCCTGGTGGTCTCCAGCTGCTGCCACACACTGGCCTTC 480  
 QY SerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGluAsn 208  
 Db 481 TCCACACAGGACTACATCAATGGCCAGATGATCTGGTATGACATCAGCAGCAGAGAAT 540  
 QY PheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuLeu 228  
 Db 541 TTTGATCGGCTTTTGGCTACGGCATAGTTCTGACATGTTCTGGCTTCTTTCCCTCCTT 600  
 QY GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248  
 Db 601 GGTCATTTTGGTGTCTATTCACTGATGGTCAGGAGCCTGATCAAGCCAGAGCAACCT 660  
 QY HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp 268  
 Db 661 CATGAGGACAGCAACACAGCCGAGCGGTCCATCGGACCATCTTACTGTTGTGTGG 720  
 QY ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeu 288  
 Db 721 CCTCTTACCCCTCTGTTTGTGGCCCTTCCATATCACTCGCTCTTCTACCTCACCATCTG 780  
 QY LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308  
 Db 781 CTTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGCCAGCAGCTGTGGCTACAGATATG 840  
 QY GluAlaSerGlyGluCysGlnGlnLeuProGlnProSerProValLeuSerPheLysGly 328  
 Db 841 GAGGCCTCTGGTGTAGTGTGAGCAGCTGCTCAACCCAGTCTCTGTACTTTCTTCAAGGG 900  
 QY GlyLysAsnArgValArgLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisPro 348  
 Db 901 GGCAAAATAGAGTCAGGCTCTCCAGAACTGAGGCAGACAAGTTGGTGAGCATCCA 960  
 QY AlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
 Db 961 GCTGGGAGAGAGATGCCAGGGTTGAACAGATCT 996

RESULT 4

AAD29667

ID AAD29667 standard; cDNA; 1076 BP.

XX

AC AAD29667;

XX

17-MAY-2002 (first entry)

Human G-protein coupled receptor (GCRC-1) cDNA.

Human; G-protein coupled receptor; GREC-1; cell proliferative disorder; neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory; metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease; Pick's disease; Huntington's disease; Parkinson's disease; hypertension; atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic; osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS; anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening; transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant; neurotropic; neuroprotective; cardiac; immunosuppressive; anorectic; virucide; gene; ss.

**Homo sapiens.**

Key	Location/Qualifiers
-----	---------------------

CDS 57..1076

/\*tag= 3

```
/product= "Human GCRC-1"
```

WO200210387-A2.

07-FEB-2002.

25-JUL-2001; 2001WO-US23433.

27-JUL-2000: 2000US-221478P.

03-AUG-2000; 2000US-223268P.

21-AUG-2000; 2000US-227054P.

08-SEP-2000; 2000US-231121P.

13-SEP-2000; 2000US-232243P.  
1E-SEP-2000; 2000US-232241P.

13-SEP-2000; 2000US-23269IP.  
22-SEP-2000: 2000US-235146P

24 JEF 2000; 200003-23J140F;

(INCY-) INCYTE GENOMICS INC.

Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR; Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walla NK; Hafalia AJA, Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L; Graul RC, Warren BA, Lee EA, Ding L;

WPI; 2002-188744/24.

P-PSDB; AAEL8640.

New human G-protein coupled receptor polypeptide for diagnosis, prevention and treatment of cell proliferative, neurological, cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic disorders -

Claim 5; Page 137-138; 150pp; English.

The invention relates to novel human G-protein coupled receptors (GCREC) and their encoding polynucleotides. GCREC is useful as an immunogen for preparing monoclonal and polyclonal antibodies. GCREC is useful for diagnosing, treating and preventing a cell proliferative disorder (e.g., hepatitis, psoriasis, cancer) a neurological disorder (e.g., epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial infarction), gastrointestinal disorder (e.g., gastritis, cirrhosis, Crohn's disease), an autoimmune/inflammatory disorder (e.g., acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma, rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity, osteoporosis), and viral infections. GCREC is useful in a number of drug screening techniques, and to analyse the proteome of a tissue or cell type. GCREC is useful for creating knockin humanised animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier or affected individuals, and as hybridization probes for mapping naturally occurring genomic sequences. GCREC is useful in Southern or

Qy 280 isSerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300  
 |||||  
 Db 897 ACTCGCTCTTCTACTCCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTTGTATG 956  
 |||||  
 Qy 300 lySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP 320  
 |||||  
 Db 957 GCACCCAGTGTGGCCTACAAGATATGGAGGCTCTGCTGAGTGTGAGCAGCTGCCTCAAC 1016  
 |||||  
 Qy 320 roSerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuGlnLysLeu 339  
 |||||  
 Db 1017 CCAGTCCCTGACTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCCAGAAACTG 1075  
 |||||  
 RESULT 5  
 AAS17747  
 ID AAS17747 standard; DNA; 850 BP.  
 XX  
 AC AAS17747;  
 XX  
 DT 26-FEB-2002 (first entry)  
 XX  
 DE Human P2Y-like G protein-coupled receptor exon 1.  
 XX  
 KW Human; ds: P2Y-like G protein-coupled receptor; GPCR; COPD;  
 KW chronic obstructive pulmonary disease; nervous system disease;  
 KW Parkinson's disease; multiple sclerosis; dementia; stroke;  
 KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;  
 KW bacterial infection; fungal infection; protozoan infection;  
 KW viral infection; pain; cancer; anorexia; bulimia; asthma; exon 1;  
 KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;  
 KW angina pectoris; myocardial infarction; ulcer; inflammation; allergy;  
 KW psychotic disorder; neurological disorder; anxiety; schizophrenia;  
 KW manic depression; delirium; severe mental retardation; dyskinesia.  
 XX  
 OS Homo sapiens.  
 XX  
 XN WO200185764-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 09-MAY-2001; 2001WO-EP05244.  
 XX  
 PR 11-MAY-2000; 2000US-203582P.  
 PR 21-FEB-2001; 2001US-269857P.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Ramakrishnan S;  
 XX  
 DR WPI; 2002-075242/10.  
 XX  
 PT New polynucleotides for producing P2Y-like G protein-coupled receptors  
 (GPCR) that are used for screening inhibitors or regulators of human  
 P2Y-like GPCR, especially useful for treating pain, cancer or  
 neurological disorders  
 XX  
 PS Claim 1; Fig 4; 114pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide encoding a P2Y-like  
 CC G protein-coupled receptor (GPCR) polypeptide, its fragment,  
 CC derivative or allele, a host cell containing an expression vector  
 CC comprising the polynucleotide and screening for agents that regulate the  
 CC GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR  
 CC polypeptide, which may be employed for screening agents that inhibit or  
 CC regulate human P2Y-like GPCR. The reagent or inhibitor of the human  
 CC P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR  
 CC disorders, particularly COPD (chronic obstructive pulmonary disease),  
 CC peripheral or central nervous system disease (e.g. Parkinson's  
 CC disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and  
 CC many other diseases and disorders listed in the specification), benign  
 CC prostatic hyperplasia or urinary incontinence. A pharmaceutical  
 CC composition containing the modulators and/or regulators of P2Y-like  
 CC GPCR is useful for modulating the activity of a P2Y-like GPCR.

CC In particular, these are useful for treating, preventing or ameliorating  
 CC infections (e.g. bacterial, fungal, protozoan or viral infections), pain,  
 CC cancer, anorexia, bulimia, asthma, acute heart failure, hypotension,  
 CC hypertension, osteoporosis, diabetes, angina pectoris, myocardial  
 CC infarction, ulcers, inflammation, allergies, psychotic or neurological  
 CC disorders (e.g. anxiety, schizophrenia, manic depression, delirium,  
 CC severe mental retardation or dyskinesias). The present sequence is  
 CC exon 1 of the P2Y-like GPCR of the invention.  
 XX

Sequence 850 BP; 165 A; 262 C; 201 G; 222 T; 0 other;

Alignment Scores:  
 Pred. No.: 1.09e-137 Length: 850  
 Score: 1541.00 Matches: 283  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 79.60% Indels: 0  
 Db: 24 Gaps: 0

US-10-023-586B-2 (1-360) x AAS17747 (1-850)

Qy 29 MetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeuCysGlnPheSerGluLys 48  
 |||||  
 Db 1 ATGGAGAGGTGGACATGAATACATCACAGGACAAAGGTCTCTGCAGTTCTCAGAGAG 50  
 |||||  
 Qy 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleLeuPheIleLeuGlyLeuProLeu 68  
 |||||  
 Db 61 TACAAGCAAGTCTACTCTCCCTGCGCTACAGTATCATCTTTATCTAGGGCTGCACATA 120  
 |||||  
 Qy 69 AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88  
 |||||  
 Db 121 AATGGCACTGTCTTGTGGCATCTCTGGGGCCAAACAAAGCGCTGGAGCTGCCACACC 180  
 |||||  
 Qy 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeuLeuProPheLeuIleIle 108  
 |||||  
 Db 181 TATCTGTGTAACCTGATGCTGGCCGACCTGCTTTATGTGCTATTCCTTCATCATC 240  
 |||||  
 Qy 109 ThrTyrSerLeuAspAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHisPhe 128  
 |||||  
 Db 241 ACCTACTCAGTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCGCAAGCTGTGCACCTC 300  
 |||||  
 Qy 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148  
 |||||  
 Db 301 CTGTTCTATATCAACCTTTACGGGACAGCATCTCTGCTGACCTGCATCTCTGTGCACCA 360  
 |||||  
 Qy 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgHisAlaTrp 168  
 |||||  
 Db 361 TTCCTAGGTGTGTCACCCACTGTGTTCGCTGCCCTACCGGACCCGAGGATGCCTGG 420  
 |||||  
 Qy 169 LeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe 188  
 |||||  
 Db 421 CTGGGACAGCAGCACCCCTGGGCGCTGGTGGTCTCCAGCTGCTGCCACACTGGCCTTC 480  
 |||||  
 Qy 189 SerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGluAsn 208  
 |||||  
 Db 481 TCCACACGGACTACATCAATGGCCAGATGATCTGATGATGACATGACCCAGCAGCAAGAA 540  
 |||||  
 Qy 209 PheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuLeu 228  
 |||||  
 Db 541 TTTGATCGGCTTTTTCCTACGGCATAGTCTTGACATGTTCTGGGCTTTCTTTCCCTCCT 600  
 |||||  
 Qy 229 GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro 248  
 |||||  
 Db 601 GGTCAATTTTGTGTGCTATTCACTGATGTCAGGAGCCCTGATCAAGCCAGAGAGAACT 660  
 |||||  
 Qy 249 HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp 268  
 |||||  
 Db 661 CATGAGGACAGCAGACAGCCCGAGCCAGGTCTCCGACCATCTACTGGTGTGTGG 720  
 |||||  
 Qy 269 ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisLeu 288  
 |||||  
 Db 721 CCTCTTCAACCTCTGTTTGTGGCCCTTCCATATCACTCGCTCTTCTACCTCACCATCTG 780  
 |||||

QY 289 LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308  
 Db 781 CTTCTCTCTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCCTTACAAGATATG 840

QY 309 GluAlaSer 311  
 Db 841 GAGGCTCT 849

# RESULT 6

AAH50975  
 ID AAH50975 standard; DNA; 510 BP.

XX AC AAH50975;

XX 28-AUG-2001 (first entry)

DE Human nGPCR12 coding sequence.

XX G protein-coupled receptor; nGPCR; seven transmembrane receptor;  
 KW signal transduction; schizophrenia; thyroid disorder; renal failure;  
 KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;  
 KW cardiovascular disease; proliferative disorder; hormonal disease;  
 KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;  
 KW attention deficit-hyperactivity disorder/attention deficit disorder;  
 KW Parkinson's disease; migraine; senile dementia; inflammatory disease;  
 KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
 KW neuroprotective; ds.

XX Homo sapiens.

XX OS

XX PN W0200136473-A2.

XX 25-MAY-2001.

XX 16-NOV-2000; 2000WO-US31581.

XX 16-NOV-1999; 9905-0165838.

XX 17-NOV-1999; 9905-0166071.

XX 19-NOV-1999; 9905-0166678.

XX 28-DEC-1999; 9905-0173396.

XX 22-FEB-2000; 2000US-0184129.

XX 28-FEB-2000; 2000US-0185421.

XX 28-FEB-2000; 2000US-0185554.

XX 02-MAR-2000; 2000US-0186530.

XX 03-MAR-2000; 2000US-0186811.

XX 09-MAR-2000; 2000US-0188114.

XX 17-MAR-2000; 2000US-0190310.

XX 21-MAR-2000; 2000US-0190800.

XX 20-APR-2000; 2000US-0198568.

XX 02-MAY-2000; 2000US-0201190.

XX 08-MAY-2000; 2000US-0203111.

XX 25-MAY-2000; 2000US-0207094.

XX (PHAA ) PHARMACIA & UPJOHN CO.

XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;

XX Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;

XX WPI; 2001-389826/41.

XX P-PSDB; AAG80935.

XX New G protein-coupled receptor (nGPCR-x) and its encoding

CC predisposition, specifically schizophrenia. nGPCRx are useful for  
 CC identifying compounds useful for treating schizophrenia. Detection of  
 CC nGPCRx in a sample is useful as a diagnostic tool for diseases or  
 CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,  
 CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular  
 CC diseases, proliferative disorders and hormonal disorders. Modulators of  
 CC nGPCRx activity have the utility for treating neurological disorders,  
 CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity  
 CC disorder/attention deficit disorder), and neuronal disorders such as  
 CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.  
 CC Additional disorders include inflammatory conditions (e.g. Crohn's  
 CC disease), rheumatoid arthritis, autoimmune disorders, cancers,  
 CC respiratory ailments such as asthma, and inflammatory diseases e.g.  
 CC inflammatory bowel disease.

SQ Sequence 510 BP; 92 A; 156 C; 126 G; 136 T; 0 other;

## Alignment Scores:

Pred. No.: 8,91e-79 Length: 510  
 Score: 921.00 Matches: 170  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 47.57% Indels: 0  
 DB: 22 Gaps: 0

US-10-023-586B-2 (1-360) x AAH50975 (1-510)

QY 83 TrpSerCysAlaThrThrTyrrLeuValAsnLeuMetValAlaAspLeuLeuTyrrValLeu 102

Db 1 TGGAGCTGTGCCACCACTATCTGGTGAACCTGTATGTGGCCGACCTGCTTTATGTGCTA 60

QY 103 LeuProPheLeuLeuThrTyrrSerLeuAspAspArgTrpPheGlyGluLeuLeu 122

Db 61 TTGCCCTTCCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 120

QY 123 CysLysLeuValHisPheLeuPheTyrrIleAsnLeuTyrrGlySerIleLeuLeuLeuThr 142

Db 121 TGCAAGCTGGTGCACCTTCCTGTTCTATATCAACCTTTACGGCAGCATCCTGCTGCTGACC 180

QY 143 CysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrrArg 162

Db 181 TGCATCTCTGTCACCACTTCCTAGGTGTGTGCCACCACTGTGCTGCTGCTGCTGCTGCTG 240

QY 163 ThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeu 182

Db 241 ACCGCGAGGATGCTGGCTGGCGCCAGCACCACTGGGCCCTGGTGGTCTCCAGCTG 300

QY 183 LeuProThrLeuAlaPheSerHisThrAspTyrrIleAsnGlyGlnMetIleTrpTyrrAsp 202

Db 301 CTGCCACACTGGGCTTCTCCACACGAGTACATCAATGGCCAGATGATCTGCTATGAC 360

QY 203 MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrrGlyIleValLeuThrLeuSer 222

Db 361 ATGACCAAGCAAGAGAAATTTGATCGGCTTTTGGCCCTACCGCATAGTCTGACATGCT 420

QY 223 GlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnLeuProAsp 242

Db 421 GGCTTTCTTCCCTTCCTGGTCAATTTGGTGTGCTATTCATGATGTCAGGAGCCTGAT 480

QY 243 GlnAlaArgGlyGluProHisGluAspArg 252

Db 481 CAAGCCAGAGAGAGAACTCATGAGGACAGG 510

## RESULT 7

AAAS30774/c

ID AAAS30774 standard; cDNA; 681 BP.

XX AC AAAS30774;

XX 04-DEC-2001 (first entry)

XX Human cDNA encoding G protein-coupled receptor nGPCR-75.

XX DE

XX

Human; G protein-coupled receptor; nGPR-x; ss; antiviral; analgesic; cytosolic; cardiant; antidiabetic; anorectic; hypotensive; hypertensive; antiparkinsonian; nootropic; neuroprotective; antidepressant; viral infection; HIV-1; human immunodeficiency virus; HIV-2; pain; cancer; metabolic disease; cardiovascular disease; type 2 diabetes; obesity; anorexia; hypotension; hypertension; myocardial infarction; atherosclerosis; Parkinson's disease; psychosis; neurological disorder; schizophrenia; migraine; major depression; anxiety; mental disorder; manic depression; dyskinesia; Huntington's disease; Tourette's Syndrome.

Homo sapiens.

WO200166750-A2.

13-SEP-2001.

08-MAR-2001; 2001WO-US07322.

08-MAR-2000; 2000US-0187581.

08-MAR-2000; 2000US-0187582.

08-MAR-2000; 2000US-0187714.

08-MAR-2000; 2000US-0187715.

08-MAR-2000; 2000US-0187825.

08-MAR-2000; 2000US-0187828.

08-MAR-2000; 2000US-0187829.

08-MAR-2000; 2000US-0187830.

08-MAR-2000; 2000US-0187833.

08-MAR-2000; 2000US-0187874.

08-MAR-2000; 2000US-0187930.

08-MAR-2000; 2000US-0188049.

08-MAR-2000; 2000US-0189294.

08-MAR-2000; 2000US-0187929.

08-MAR-2000; 2000US-0187928.

(PHAA ) PHARMACIA & UPJOHN CO.

Vogeli G, Wood LS;

WPI; 2001-536778/59.

P-PSDB; AAU19205.

Isolated nucleic acid molecules encoding G protein-coupled receptors termed nGPR-x, useful in the treatment and diagnosis of viral infections, cancers and mental disorders (e.g. Parkinson's disease and schizophrenia) -

Claim 4; Page 198; 336pp; English.

The invention relates to novel isolated nucleic acid molecules encoding G protein-coupled receptors termed nGPR-x. nGPR-x polynucleotides, polypeptides, and modulators may be used in the treatment of diseases and conditions such as infections, such as viral infections caused by HIV-1 (human immunodeficiency virus) or HIV-2, pain, cancers, metabolic and cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity, anorexia, hypotension, hypertension, myocardial infarction, atherosclerosis), Parkinson's disease, and psychotic and neurological disorders, including schizophrenia, migraine, major depression, anxiety, mental disorder, manic depression, and dyskinesias, such as Huntington's disease or Tourette's Syndrome and many other diseases and syndromes listed in the specification. nGPR-x polynucleotides and polypeptides, as well as nGPR-x modulators, may also be used in diagnostic assays for such diseases or conditions. The present sequence encodes a G protein-coupled receptor of the invention.

Sequence 681 BP; 163 A; 171 C; 171 G; 176 T; 0 other;

Alignment Scores:

Pred. No.: 1.53e-60 Length: 681  
Score: 732.00 Matches: 132  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 37.81% Indels: 0

DB:	22	Gaps:	0
US-10-023-586B-2 (1-360) x AAS30774 (1-681)			
Qy	229	GlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGluPro	248
Db	681	GGTCACTTTGGTGGTCTATTCACTGATGGTCAGGAGCCTGATCAAGCCAGAGGAGAACCT	622
Qy	249	HisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrp	268
Db	621	CATGAGGACAGCAACACAGCCGAGCCAGGTCATCCGGACCATCTCTACTGGTGTGG	562
Qy	269	ProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeu	288
Db	561	CCTCTTCACTCTCTGTTTGTGGCTTCCATATCACTACCTCGCTCTTCTACCTACCATCTG	502
Qy	289	LeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet	308
Db	501	CTTCTGCTTCTCAGGACTGCCAGCTCTTATGACGCCAGCTGGCTTCAAGATATG	442
Qy	309	GluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPheLysGly	328
Db	441	GAGGCTCTGCTGAGTGTGAGCAGCTGCTCAACCCAGTCTCTGTACTTTCTTCAAGGG	382
Qy	329	GlyLysAsnArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisPro	348
Db	381	GGCAAAATAGAGTCAGCTCTCTCCAGAACTGAGGCGAGAACAGTTGGTGAGCATCCA	322
Qy	349	AlaGlyArgLysArgCysProGlyLeuAsnArgSer	360
Db	321	GCTGGGAGAGAGATGCCAGGGTTGAACAGATCT	286
RESULT 8			
AAS30783			
ID	AAS30783	standard; cDNA; 585 BP.	
XX	AC	AAS30783;	
XX	DT	04-DEC-2001 (first entry)	
XX	DE	Human cDNA encoding G protein-coupled receptor nGPR-84.	
XX	KW	Human; G protein-coupled receptor; nGPR-x; ss; antiviral; analgesic; cytosolic; cardiant; antidiabetic; anorectic; hypotensive; hypertensive; antiparkinsonian; nootropic; neuroprotective; antidepressant;	
XX	KW	viral infection; HIV-1; human immunodeficiency virus; HIV-2; pain; cancer; metabolic disease; cardiovascular disease; type 2 diabetes; obesity; anorexia; hypotension; hypertension; myocardial infarction; atherosclerosis; Parkinson's disease; psychosis; neurological disorder; schizophrenia; migraine; major depression; anxiety; mental disorder; manic depression; dyskinesia; Huntington's disease; Tourette's Syndrome.	
XX	OS	Homo sapiens.	
XX	PN	WO200166750-A2.	
XX	PD	13-SEP-2001.	
XX	PF	08-MAR-2001; 2001WO-US07322.	
XX	PR	08-MAR-2000; 2000US-0187581.	
XX	PR	08-MAR-2000; 2000US-0187582.	
XX	PR	08-MAR-2000; 2000US-0187714.	
XX	PR	08-MAR-2000; 2000US-0187715.	
XX	PR	08-MAR-2000; 2000US-0187825.	
XX	PR	08-MAR-2000; 2000US-0187828.	
XX	PR	08-MAR-2000; 2000US-0187829.	
XX	PR	08-MAR-2000; 2000US-0187830.	
XX	PR	08-MAR-2000; 2000US-0187833.	
XX	PR	08-MAR-2000; 2000US-0187874.	
XX	PR	08-MAR-2000; 2000US-0187930.	
XX	PR	08-MAR-2000; 2000US-0188049.	
XX	PR	08-MAR-2000; 2000US-0189294.	



08-MAR-2000; 2000US-0187929.  
08-MAR-2000; 2000US-0187928.

(PHAA ) PHARMACIA & UPJOHN CO.

Vogeli G, Wood LS;

WPI; 2001-536778/59.

P-PSDB; AAU19214.

Isolated nucleic acid molecules encoding G protein-coupled receptors termed nGPCR-x, useful in the treatment and diagnosis of viral infections, cancers and mental disorders (e.g. Parkinson's disease and schizophrenia) -

Claim 4; Page 201; 336pp; English.

The invention relates to novel isolated nucleic acid molecules encoding G protein-coupled receptors termed nGPCR-x. nGPCR-x polynucleotides, polypeptides, and modulators may be used in the treatment of diseases and conditions such as infections, such as viral infections caused by HIV-1 (human immunodeficiency virus) or HIV-2, pain, cancers, metabolic and cardiovascular diseases and disorders (e.g., type 2 diabetes, obesity, anorexia, hypotension, hypertension, myocardial infarction, atherosclerosis), Parkinson's disease, and psychotic and neurological disorders, including schizophrenia, migraine, major depression, anxiety, mental disorder, manic depression, and dyskinesias, such as Huntington's disease or Tourette's Syndrome and many other diseases and syndromes listed in the specification. nGPCR-x polynucleotides and polypeptides, as well as nGPCR-x modulators, may also be used in diagnostic assays for such diseases or conditions. The present sequence encodes a G protein-coupled receptor of the invention.

Sequence 585 BP; 145 A; 142 C; 155 G; 143 T; 0 other;

#### Alignment Scores:

Pred. No.: 9.19e-56 Length: 585  
Score: 681.00 Matches: 131  
Percent Similarity: 96.32% Conservative: 0  
Best Local Similarity: 96.32% Mismatches: 2  
Query Match: 35.18% Indels: 3  
Gaps: 0

US-10-023-586B-2 (1-360) x AAS30783 (1-585)

QY 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGlu 247  
DB 3 CTTGGTCAATTTTGGTGTGCTATTACTGATGGTCAGAGGCTGATCAAGCCAGAGGAGTA 62  
QY 248 -ProHisGlu-AspArgGlnHisSerProSerGlnValHisPro-AspHisProThrGly 266  
DB 63 ACCTCATGAGTACAGGCACACAGCCGAGCAGGTCCATCCGGGACCATCTACTGGT 122  
QY 267 ValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHis 286  
DB 123 GTCTGGCCCTCTCACCTCTGTTTGGCCCTCCATATCATCTGCTCTTCTACCTCAC 182  
QY 287 HisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGln 306  
DB 183 CATCTGCTTCTCTTCTCAGACATGCCACCTCTTGTATGGAGCCAGTGTGCCCTACAA 242  
QY 307 AspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPhe 326  
DB 243 GATATGAGGCGCTCTGGTGTAGTGTGACAGCTGCCTCAACCCAGTCTGCTGACTTCTTC 302  
QY 327 LysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGlu 346  
DB 303 AAGGGGGGCAAAATAGAGTCAGCTCTCTCCAGAACTGAGGAGGAGCAACAGTTGGGTGAG 362  
QY 347 HisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
DB 363 CATCCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT 404

#### RESULT 9

AA04981

ID AAD04981 standard; cDNA; 1163 BP.

XX AC AAD04981;

XX DT 17-JUL-2001 (first entry)

XX DE Human purinergic receptor P2Y2 cDNA.

XX KW Human; purinergic receptor; P2Y2; antidepressant; vulnerable; hypotensive;  
KW anticonvulsant; antiinflammatory; therapy; nervous system disorder;  
KW epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;  
KW irritable bowel disorder; reproductive system disorder; hypertension;  
KW peripheral vascular disease; immune system disorder; chronic bronchitis;  
KW premature ejaculation; asthma; neuromuscular disease; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
FT primer\_bind complement (1..32)  
FT /tag= a  
FT /note= "Sense primer"

FT CDS 15..1148  
FT /tag= b  
FT /product= "Protein encoded by human P2Y2 cDNA"

FT /transl\_except= (pos:786..788, aa:Phe)  
FT 787  
FT misc\_feature  
FT /tag= c  
FT /note= "Represented in the specification as F"

FT primer\_bind 1131..1163  
FT /tag= d  
FT /bound\_moiety= "Antisense primer"

XX US6214581-B1.

XX PD 10-APR-2001.

XX PF 13-NOV-1998; 98US-0191136.

XX PR 16-JAN-1998; 98US-0071298.

XX PR 16-JAN-1998; 98US-0071569.

XX PR 16-JAN-1998; 98US-0008185.

XX PR 16-JAN-1998; 98US-0008526.

XX PA (ABBO ) ABBOTT LAB.

XX PI Lynch KJ, Burgard EC, Van Biesen T;

XX DR WPI; 2001-315459/33.

XX DR P-PSDB; AAE01144.

XX PT Novel isolated polynucleotide encoding human purinergic P2X3 receptor polypeptide useful for identifying potentially therapeutic compounds that modulate or otherwise interact with P2X containing receptors -

XX PS Example 14; Fig 12A; 53pp; English.

XX CC The present sequence is a cDNA encoding human P2Y2 receptor protein. P2X receptors are ligand-gated ion channels while P2Y receptors operate generally through a G-protein coupled system. P2X purinoreceptor drugs are potential therapeutic agents in several disorders including central nervous system or peripheral nervous system conditions, e.g., epilepsy, pain, depression, neurodegenerative disorders, disorders of the skeletal muscle such as neuromuscular diseases, disorders of the reproductive system, asthma, peripheral vascular disease, hypertension, immune system disorders, irritable bowel disorder, premature ejaculation, cystic fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the activity of extracellular nucleotide triphosphates to regulate chloride secretion in human airway epithelia.

XX CC Sequence 1163 BP; 177 A; 403 C; 332 G; 250 T; 1 other;

Alignment Scores:  
Pred. No.: 1.49e-33 Length: 1163  
Score: 453.00 Matches: 129  
Percent Similarity: 51.37% Conservative: 40  
Best Local Similarity: 39.21% Mismatches: 138  
Query Match: 23.40% Indels: 23  
DB: Gaps: 10

US-10-023-586B-2 (1-360) x AAD04981 (1-1163)

Qy 43 CysGlnPheSerGluLysTyrlValLeuSerLeuAlaTyrSerIleIlePhe 62  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 87 TGCCGGCTTCACAGGACTTCAAGTACGTGTGCCTGCTGCTACGGCGTGTTGCC 146  
Qy 63 IleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArg 82  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 147 GTGCTGGCGTGTCTGAACCGCTGGCGCTCATATCTTCTTGTGGCGCTCAGACC 206  
Qy 83 TrpSerCysAlaThrThrTyrlValLeuValAsnLeuMetValAlaAspLeuValLeu 102  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 207 TGGAACTGGTCCACACATATATTCCACCTGGCTGTGTGTGTGTGTGTGTGTGTGT 266  
Qy 103 --LeuProPheLeuIleThrTyrlSerLeuAspArgTrpProPheGlyGluLeu 121  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 267 TCCCTGGCGTGTGTGTATTAATACCGCGCGCGCACCTGGCCCTTCAGACGGTG 326  
Qy 122 LeuCysLysLeuValHisPheLeuPheTyrlIleAsnLeuTyrlGlySerIleLeuLeu 141  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 327 CTCGTCAAGCTGGTGGCTTCTTCTTACACCACTTTTACTGACGATCTCTTCTC 386  
Qy 142 ThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyr 161  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 387 ACCTGCATCGCTGCACCGTGTCTGGCGCTTACGACCTCTGCGCTCCCTGGCGTG 446  
Qy 162 ArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGln 181  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 447 GGCGGGCGCGCTACGCTCGCGGGTGGCGGGCGCGGTGGTGGTGGTGGTGGTGG 506  
Qy 182 LeuLeuProThrLeuAlaPheSerHisThrAspTyrlIleAsnGlyGlnMetIleTrpTyr 201  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 507 CAGCCCCCGTGTCTACTTGTCCACACAGCGCGCGGGCGCGGTAACTGCGCAC 566  
Qy 202 AspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyr---GlyIleValLeuThr 220  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 567 GACACCTCGCACCGGACTCTTCAGCGCTTCGTGGCTCAGCTCAGTCTGCTGGCG 626  
Qy 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 627 CTGCTCTTCGCGGT-GCCCTTTCGCGTATCTTCTGTTCAGTGTCTGCTGGCG 685  
Qy 241 ProAspGlnAlaArg-----GlyGluProHisGluAspArgGlnHisSerPro 256  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 586 ACTGCTAAGCCACCTACGGGACTCTGGGCGGCT-----GCCTAGGGCCAA 733  
Qy 257 SerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAla 276  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 734 CGCAAGTCCGTGGCGCACCATCGCGTGTGTCTGCTGCTGCTGCTGCTNCCCTGCC 793  
Qy 277 LeuProTyrHisSerLeuLeuLeuProHisHis---LeuLeuSerAlaPheSerGly--- 294  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 794 ATTCACAGTCA-----CCGACACCTCTACTACTCTCTCTCTCTCTCTCTCTCT 841  
Qy 295 --LeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGlu 313  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 842 CAGCTGCCACACCTCAACGGCCATACATACATACATACATACATACATACATACAT 901  
Qy 314 CysGluGlnLeuProGlnPro-SerProValLeuSerPheLysGlyLysAsnArgVa 333  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 902 TGCTAACAGTTG-----CCTTGACCGCTGCTCTACTTCTCTCTCTCTCTCTCTCT 955  
Qy 333 lArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlnHisProAlaGlyArgLysAr 353  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```
Db 220 CTCAGC-----CCAGGTCTCCGGCAGCAGTGCAGTGGAGCTGGACTGTGG 264
Qy 45 PheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeu 64
Db 265 TTTGATGAGATTCAAGTTTCATCTGCTGCCCTGTGAGCTATGCAGTTGCTTTGGCTG 324
Qy 65 GlyLeuProLeuAsnGlyThrValLeuThrPheHisSerTrpGlyGlnThrLysArgTrpSer 84
Db 325 GGCTTGGGCTTAAGCGCCCAACCCATGCTCTTCATCTTCGCGCTCCGACCCCTGGAT 384
Qy 85 CysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---Leu 103
Db 385 GCAACGCCCACTACATGTTCCACCTGGCATCTGCACACCTGTATGTGCTGCTGCTG 444
Qy 104 ProPheLeuIleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCys 123
Db 445 CCACCTCATCTACTATTATGAGCCCAACCACTGGCCCTTTGGCACTGAGACTGTC 504
Qy 124 LysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCys 143
Db 505 AAGTTCGCTCGCTTCTTTCTATTGAACTCTACTGCAGTGCCTTTTCCCTCACCTGC 564
Qy 144 IleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThr 163
Db 565 ATCAGCGTGCACCGCTACCTGGGCATCTGCCACCCACTTCGGGCACCTACCTGGGCCGC 624
Qy 164 ArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuLeuLeu 183
Db 625 CCTCGCTCCGAGCCTTCTCGCTGGCAGTTTGGTTGGTGTGCTAGCCGCTGCCCTGCTG 684
Qy 184 ProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMet 203
Db 685 CCCAACCTGTCTTGTGCACACCAAGAGGACCACCGCTCTGCTGCCATGCACACC 744
Qy 204 ThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGly 223
Db 745 ACTCGGCTGAAGATTTTGACCA-CTATGTGCA-----CCTCAGCTCGCGCT 791
Qy 224 -----PheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGly 238
Db 792 CATGGGCTGCTCTTTGGCGTGCCTGCTGCTCACTCTTGTGTGCTATGGACTCATGGC 851
Qy 239 GlnGluProAspGlnAlaArgGly-----GluProHisGluAsp 251
Db 852 TCGTGGCTGTATCAGCGCTTGGCAGGCTCTGCACAGCTCTCTTCTCGCT----- 902
Qy 252 ArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrpProLeuHis 271
Db 903 -----CCGCTCTCCGCACCATAGCTGTGGTGTGCTGCTGCTGTCTTTC 944
Qy 272 ProLeuPheCysAlaLeuProTyrHisSerLeuLeuLeuProHisHisLeuLeu----- 289
Db 945 TGTGCTGCTGCTGCTTTCACATCAC-----CCGCACCATTTTACCTACCTGGC 992
Qy 290 ---SerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMet 308
Db 993 CAGGCTTTGGAAGCTGACTCGCGAGTACTGAACATTGTCAACGTGCTATATAAGTGC 1052
Qy 309 GluAlaSerGlyGluCysGluGlnLeuProGlnProSerProValLeuSerPheLysGly 328
Db 1053 TCGGCCCTGGCCAGTCCCAACAGCTGCCGTGATCCTGTGCTCTACTTGTCTACTGGGA 1112
RESULT 11
AAD04980
ID AAD04980 standard; DNA: 2025 BP.
XX
AC AAD04980;
XX
XX
DT 17-JUL-2001 (first entry)
XX
DE Human purinergic receptor P2Y2 DNA.
XX
KW Human; purinergic receptor; P2Y2; antidepressant; vulnery; hypotensive;
```

```
KW anticonvulsant; antiinflammatory; therapy; nervous system disorder;
KW epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;
KW irritable bowel disorder; reproductive system disorder; hypertension;
KW peripheral vascular disease; immune system disorder; chronic bronchitis;
KW premature ejaculation; asthma; neuromuscular disease; ds.
XX
OS Homo sapiens.
FH Key Location/Qualifiers
CDS 246..1379
FT /tag= a
FT /product= "Human P2Y2 receptor protein"
FT primer_bind complement (246..263)
FT /tag= b
FT /bound_moiety= "Sense primer"
FT primer_bind 1362..1379
FT /tag= c
FT /bound_moiety= "Antisense primer"
FT
FT US6214581-B1.
PN
XX 10-APR-2001.
XX
XX 13-NOV-1998; 98US-0191136.
XX
XX 16-JAN-1998; 98US-00711298.
PR 16-JAN-1998; 98US-00711669.
PR 16-JAN-1998; 98US-0008185.
PR 16-JAN-1998; 98US-0008526.
XX
PA (ABBO ) ABBOTT LAB.
XX
XX Lynch KJ, Burgard EC, Van Biesen T;
XX
DR WPI: 2001-315459/33.
DR P-PSDB; AAE01143.
XX
PT Novel isolated polynucleotide encoding human purinergic P2X3 receptor
PT polypeptide useful for identifying potentially therapeutic compounds
PT that modulate or otherwise interact with P2X containing receptors -
XX
PS Example 14; Fig 11A; 53pp; English.
XX
CC The present sequence is a DNA encoding human wild type P2X4 receptor
CC protein. P2X receptors are ligand-gated ion channels while P2Y receptors
CC operate generally through a G-protein coupled system. P2X purinoreceptor
CC drugs are potential therapeutic agents in several disorders including
CC central nervous system or peripheral nervous system conditions, e.g.,
CC epilepsy, pain, depression, neurodegenerative disorders, disorders of the
CC skeletal muscle such as neuromuscular diseases, disorders of the
CC reproductive system, asthma, peripheral vascular disease, hypertension,
CC immune system disorders, irritable bowel disorder, premature ejaculation,
CC cystic fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the
CC activity of extracellular nucleotide triphosphates to regulate chloride
CC secretion in human airway epithelia.
XX
SQ Sequence 2025 BP; 388 A; 617 C; 595 G; 425 T; 0 other;
```

## Alignment Scores:

Pred. No.:	3,69e-32	Length:	2025
Score:	442.00	Matches:	128
Percent Similarity:	50.91%	Conservative:	40
Best Local Similarity:	38.79%	Mismatches:	138
Query Match:	22.83%	Indels:	25
DB:	22	Gaps:	10

US-10-023-586B-2 (1-360) x AAD04980 (1-2025)

```
Qy 43 CysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePhe 62
Db 318 TGGCGCTTCACGAGGACTTCAAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377
Qy 63 IleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheHisSerTrpGlyGlnThrLysArg 82
```



```
QY 104 ProPheLeuLeuIleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCys 123
Db 238 CCCTACTTATCTATAACTACGCAGAGGGACCACCTGGCCCTTCGAGACCTCGCGTGC 297
QY 124 LysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCys 143
Db 298 CGCTTTGTAGCTCTCTTCTATGCAATCTACATGGCAGCATCCTGTTCCTCAGCTGC 357
QY 144 IleSerValHisPheLeuValCysHisProLeuLeuCysSerLeuProTyrArg-- 162
Db 358 ATTAGCTTCCAGCGCTACCTGGCATCTGCCACCCCTGGCTTCCTGGCACAAGCGTGA 417
QY 163 ThrArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeu 182
Db 418 GGTCCGCGTCTGCTGGTGGTAGTGTGGAGTGTGTGGCTGTGTGACAGCCAGTGC 477
QY 183 LeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAsp 202
Db 478 CTGCCACGGCAGTCTTTGTGTCACAGGATCCAGCGCAACCGCACTGTGTGCTACGAC 537
QY 203 MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSer 222
Db 538 CTGAGCCCACTCCTGTCTACTCGCTACCTGCCCTATGGTATGGCCCTCACGGTCATC 597
QY 223 GlyPheLeu--SerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluPro 242
Db 598 GGCTTCTGTGCTGCCCTCATAGCCTTACTGGCTGTATTGTGCG----- 641
QY 242 spGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHisPro 262
Db 642 -----CATGCCCGCGCCCTGTGTGCCAG 666
QY 262 spHisProThr----- 265
Db 667 GATGGCCACAGCTCCTGTGGCCCAAGAGCGCGCAGCGGCTCGTATGGCTGTG 726
QY 266 -----GlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeu 284
Db 727 GTGGTGGCAGCTCTCTTGTGCATCAGCTT---CCTGGCTTTCCACATCAACAGAGAGCC 783
QY 284 euProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysG 304
Db 784 TACTTGGCTGTGGCTCCAGCGCGGTCTCTTGGCTGTGTGGCTGTGGAGACCTTCGCTGCT 843
QY 304 lyLeuGlnAspMetGluAlaSerGlyGluCysGluGln----- 316
Db 844 GCCTACAAGGACCTCGGCCCTTCGCCAGGTCAACAGTCTTCTGGACCCCATCTCTCTTC 903
QY 317 -----LeuProGlnProSerProValLeuSerPheLysGlyGlyLysA 331
Db 904 TACTTCACACACAGAGATTCCGGCGGC-AACCCACAGATCTCTTACAGAGGCTCAGAG- 961
QY 331 snArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisProAlaGlyA 351
Db 962 -----CCAAGTGGCA 971
QY 351 rglYsArg 353
Db 972 GAGGCAGA 979
RESULT 13
ID AAQ88134
XX AAQ88134 standard; cDNA; 1842 BP.
AC AAQ88134;
XX
XX
XX
XX 29-NOV-1995 (first entry)
DT Human P20 receptor gene.
DE Epithelial mucosa; mucus; cystic fibrosis; asthma; ss;
XX
XX
XX
XX
```

```
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 57..1181
FT /*tag= a
XX
XX W09510538-A.
XX
XX 20-APR-1995.
XX
XX 04-OCT-1994; 94WO-US11260.
XX
XX 15-OCT-1993; 93US-0138137.
XX (UMOR ) UNIV MISSOURI.
XX (UYNC-) UNIV NORTH CAROLINA.
XX
XX Boucher RC, Erb LJ, Harden TK, Lustig KD, Parr CE;
XX Sullivan DM, Turner JT, Weisman GA;
XX
XX WPI; 1995-169967/22.
XX P-PSDB; AAR72457.
XX
XX DNA encoding human P2U receptor and null cells expressing the
XX receptors - for stimulating or inhibiting growth of cultures of
XX mammalian cells, and for treating diseases of airway epithelial
XX mucosa, e.g. asthma.
XX
XX Claim 3; Page 34; 47pp; English.
XX
XX The sequence is that of the human P20 receptor gene. The gene may
XX be used to treat diseases of epithelial mucosal surfaces by
XX enhancing or inhibiting mucus prodn. The gene may be used to treat
XX cystic fibrosis, asthma and chronic bronchitis.
XX See also AAQ88135-40.
XX
XX SQ Sequence 1842 BP; 364 A; 557 C; 528 G; 393 T; 0 other;

Alignment Scores:
Pred. No.: 9.82e-30 Length: 1842
Score: 416.00 Matches: 127
Percent Similarity: 49.85% Conservative: 37
Best Local Similarity: 38.60% Mismatches: 141
Query Match: 21.49% Indels: 25
DB: 16 Gaps: 11

US-10-023-586B-2 (1-360) x AAQ88134 (1-1842)
QY 43 CysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePhe 62
Db 129 TGGCGGTTCAACGAGGACTTGAAGTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 188
QY 63 IleLeuGlyLeuProLeuAsnGlyThrValLeuThrHisSerTrpGlyGlnThrLysArg 82
Db 189 GTCTTGGGCTGTGCTGAACGCGGCTCTACATCTTCTGTGCGCCCTCAAGACC 248
QY 83 TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaLeuTyrValLeu 102
Db 249 TGAATCGGTCCACACATATATGTTCCACCTGGCTGTGTCTGATGACATGTATGCGGCC 308
QY 103 ---LeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGluLeu 121
Db 309 TCCCTGCCGCTGCTGTGCTATTACTACGCCCGCGGCGACCATCGCCCTTCAGCAGGTG 368
QY 122 LeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeu 141
Db 369 CTCTGCAAGCTGGTGGCTTCTCTCTACCAACCTTTACTGACATCTCTTCTCTCTC 428
QY 142 ThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyr 161
Db 429 ACCTGATCAGCGTGCAGCGGTGTCTGGGCGTCTTACGACCTTCCTGCGCTCCCTGCG 488
QY 162 ArgThrArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGln 181
```



```
Db 342 TGTGCTTCAACGAGGACTTCAAGTAGCTGTGTGCTGCTGCTATGCGGTGGTGC 401
Qy 63 ILeuGlyLeuProLeuAsnGlyThrValLeuThrHisSerTrpGlyGlnThrLysArg 82
Db 402 GTGCTGGGTGGCTGACGCGTGGCTCTATATCTTCCTATGCGCCCTCAAAACC 461
Qy 83 TrpSerCysAlaThrThrThrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu 102
Db 462 TGGAAACGCTCCACCACTACATGTTTCACCTGGCAGTTTCGGACTCTCTCTACGACGG 521
Qy 103 --LeuProPheLeuLeuLeuThrThrSerLeuAspAspArgTrpPheGlyGluLeu 121
Db 522 TCCCTGCGCTGTGGTGTATTAATCTAGCCCGGGGTGACCACTGGCCATTATGACCGGTG 581
Qy 122 LeuCysLysLeuValHisPheLeuPheThrIleAsnLeuTyrGlySerIleLeuLeuLeu 141
Db 582 CTCTGCAAGCTGGTGGCTTCTCTCTACACCACTCTACTGCACGATCCTCTCCCTC 641
Qy 142 ThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyr 161
Db 642 ACCTGCATCAGCGTGCACCGGTGCTGGAGTCTCGCGCCTCTGCACTCCCTGCGTGG 701
Qy 162 ArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGln 181
Db 702 GCGCGGCGCGCTTATGCGCGGGTGGCTGCGGTGTGTGGTGTGGTGTGGTGTGGTGC 761
Qy 182 LeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyr 201
Db 762 CAGCACCGCTGCTCTACTTCTGTCACCACTGCGGTGGGGAACCCGGATCACTTGCAT 821
Qy 202 AspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeu 221
Db 822 GACACCTCGCGCGGAGAGCTCTTAGCATTCTGTGGCTTACAGCTCCGTCATGCTGGG- 880
Qy 222 SerGlyPhe--LeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
Db 881 TCTGCTTTTGTGCTGCGCTTTTCCGTAATCCTGCTGTGTGCTGTATGCGCCAGGCG 940
Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
Db 941 GCTGCTCAACCGCTTATGGACCAACGAGGCTCTGCTCGGCGCAAGCGCAAGTCTGT 1000
Qy 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
Db 1001 GCGCACCATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1060
Qy 281 Ser-----LeuLeuLeuProHisHisLeuSerAlaPheSerGlyLeuProAla 297
Db 1061 GCGCACCTCTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1108
Qy 298 LeuAspGlySerGlyGlyLeuGlnAspMetGluAlaSerGlyGlyCysGluGlnLeu 317
Db 1109 CCTCAAGCCCATCAACATGGCATATAGATCAACCCGCGCTGCGCCAGCCCAACAGTTG 1168
Qy 318 ProGlnPro-SerProValLeuSerPheLysGlyGlyAsnArgValArgLeuGln 337
Db 1169 -----TCTTGACCGGCTACTTCTCTGCGGCGCTGCTGCTGCTGCTGCTGCTGCTG 1222
Qy 337 nLysLeuArg-----GlnAsnLysLeuGlyGln 346
Db 1223 AGATGCCAACCCACCGAGGAGCTACCCCGCCACGAGCTCGTGGCAAGCTGGGCGCT 1282
Qy 346 uHisProAlaGlyArg 351
Db 1283 GCACAGGCTTAACAGA 1298
RESULT 15
ID AAD32944
XX AAD32944 standard; DNA: 2138 BP.
AC AAD32944;
XX
XX 01-JUL-2002 (first entry)
```

```
XX Mus musculus GPCR ATP-receptor P2U1 gene targeting construct.
DE Mouse; G-protein coupled receptor; GPCR; transgenic animal; receptor;
KW genetic therapy; pharmacological; genetic disease; neuropsychological;
KW neurological; psychotic illness; nephrotropic; gynaecological;
KW psychostimulant; ATP-receptor P2U1; ds.
XX Mus musculus.
XX Key Location/Qualifiers
FT misc_feature 286..350
FT /tag= a
FT /note= "Sequence flanking Neo insert in targeting
FT misc_feature 351..636
FT /tag= b
FT /note= "Sequence deleted in targeting construct"
FT misc_feature 637..941
FT /tag= c
FT /note= "Sequence flanking Neo insert in targeting
FT construct"
XX WO200203793-A2.
XX 17-JAN-2002.
XX 10-JUL-2001; 2001WO-US21923.
XX 10-JUL-2000; 2000US-217058P.
XX 10-JUL-2000; 2000US-217179P.
XX 10-JUL-2000; 2000US-217223P.
XX 10-JUL-2000; 2000US-217253P.
XX 10-JUL-2000; 2000US-217255P.
XX 10-JUL-2000; 2000US-217256P.
XX 10-JUL-2000; 2000US-217257P.
XX 11-JUL-2000; 2000US-217347P.
XX 12-JUL-2000; 2000US-217537P.
XX 12-JUL-2000; 2000US-218069P.
XX 12-JUL-2000; 2000US-218074P.
XX 12-JUL-2000; 2000US-218358P.
XX 27-AUG-2000; 2000US-221483P.
XX 07-AUG-2000; 2000US-223120P.
XX 26-OCT-2000; 2000US-223122P.
XX 15-NOV-2000; 2000US-243958P.
XX 20-NOV-2000; 2000US-249408P.
XX 16-JAN-2001; 2001US-262113P.
XX 16-JAN-2001; 2001US-262205P.
XX (DELT-) DELTAGEN INC.
XX Allen KD, Brennan TJ;
XX WPI; 2002-164574/21.
XX Novel non-human transgenic animal, especially transgenic mice useful
XX for identifying an agent that modulates expression or function of
XX target gene, comprises disruptions in target G protein-coupled receptor
XX gene.
XX Disclosure; Fig 20A; 101pp; English.
XX The invention relates to a non-human transgenic animal having targeted
XX G-protein coupled receptor (GPCR) gene disruptions in melanocortin-3 gene
XX 5-HT-2B gene, chemokine receptor 9A gene, glucocorticoid-induced receptor
XX gene, orphan GPR10 (UHR)-1 gene, orphan GPR14 gene, orphan GPR15 gene,
XX beta chemokine receptor (E01) gene, endothelial differentiation GPCR3
XX (EDG3) gene, ATP receptor P2U1 gene or adenosine 3 receptor gene. The
XX transgenic animal is useful for identifying an agent that modulates the
XX expression or function of the target gene, for identifying an agent that
XX ameliorates a phenotype associated with a disruption in the target gene.
```

CC A transgenic construct is useful for producing a transgenic animal,  
CC preferably a transgenic mouse. The transgenic animal is useful for  
CC testing the efficacy of proposed genetic and pharmacological therapies  
CC for human genetic diseases, such as neurological, neuropsychological or  
CC psychotic illnesses. The transgenic animal is also useful as models for  
CC diseases, disorders or conditions associated with phenotypes relating to  
CC a disruption in a target, and to identify pharmaceuticals, therapies,  
CC drugs and interventions which may be effective in treating a disease or  
CC other phenotypic characteristics of the animal. An agent which modulates  
CC the expression of the target gene is useful as a therapeutic for treating  
CC conditions associated with a disruption of the target gene. The present  
CC sequence is mouse GPCR ATP-receptor P2U1 gene targeting construct.  
XX

SQ Sequence 2138 BP; 423 A; 623 C; 608 G; 480 T; 4 other;

Alignment Scores:  
Pred. No.: 1.51e-29 Length: 2138  
Score: 415.00 Matches: 119  
Percent Similarity: 49.69% Conservative: 43  
Best Local Similarity: 36.50% Mismatches: 141  
Query Match: 21.44% Indels: 24  
DB: 24 Gaps: 6

US-10-023-586B-2 (1-360) x AAD32944 (1-2138)

43 CysGlnPheSerGluLysTyrGlnValTyrSerLeuAlaTyrSerIleIlePhe 62  
Db |||:|||||:||||| ||| ||| |||:||||| |||:|||||  
342 TGTCTGCTCAACGAGGACTTCAAGTACGTGCTGTTGGCCGCTCTATGGCGTGGTGTGC 401  
Qy 63 IleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThrIysArg 82  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
402 GTGCTGGGGTTGTGCTGAACGCTGCTGCTCTATATCTCTCTATGTCCTGAGCCGCTCAAAACC 461  
Qy 83 TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTrpValLeu 102  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
462 TGGACGGCTCCACACCTACATGTTTCACTGGGAGTTCGGACTCTCTACGACGCG 521  
Qy 103 ---LeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGluLeu 121  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
522 TCCCTGCGCGCTGTTGTTTATTACTACCGCGGGGTGACCACTGGCCATTTAGCAGCGTG 581  
Qy 122 LeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeu 141  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
582 CTCGCAAGCTGGTGGCTTCTCTCTACACCACTCTACTGCGACGATCCTCTCTCTC 641  
Qy 142 ThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyr 161  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
642 ACCTCATCAGCGTCACCGGTGCTGGGAGTCTGCGCCCTCTGCACTCCCTGCGTTGG 701  
Qy 162 ArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGln 181  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
702 GCGCGGGCGCTTATGCGCGCGGTGCTGGGTGTTGTGGTGTGGTGTGGTGTGGTGTGC 761  
182 LeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyr 201  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
762 CAGGACCCCGTCTCTCTCTACACCACTGCGGGGAGCCCGGATCATTGCGCCAT 821  
Qy 202 AspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeu 221  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
822 GACACCTCGGCCGAGAGCTCTTTAGCCATTTTGGGCTTACAGCTCCGTCATGCTGGG- 880  
Qy 222 SerGlyPhe---LeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
881 TCTGCTTTTGTGTGGCTTTTCCGTAATCTGCTGTGTACGTGCTTATGGCCAGGCG 940  
Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
941 GCTGCTCAACCGGCTTATGGGACACAGAGGCTGCTCGGGCCAAAGCGCAAGCTCTGT 1000  
Qy 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
1001 CGCACCATTGCTGGTACTGCGCGTCTTCGCCCTCTCTGCTTTCTGCTTTCCAGCTCAC 1060

Qy 281 Ser-----LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAla 297  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
1061 GCGCACCCTCTACTACTCTCTTCCG-----ATCACTTGACCTCAGCTGCCACAC 1108  
Qy 298 LeuAspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeu 317  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
1109 CCTCAACGCCCATCAACATGGCATATAGATCATCCCGCGCTGGCCGACGACAGTTG 1168  
Qy 318 ProGlnPro-SerProValLeuSerPheLysGlyLysAsnArgValArgLeuLeuG1 337  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
1169 -----TCTTGACCGGTACTCTACTTCTGCGAGGCGAGAGACTTGTCCGCTTTGCCG 1222  
Qy 337 nLysLeuArg-----GlnAsnLysLeuGlyG1 346  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
1223 AGATGCCAAAGCCACCACGGAGCCTACCCCGACGACGCTCGTCGCAAGCTGGGCT 1282  
Qy 346 uHisProAlaGlyArg 351  
Db |||:|||||:||||| |||:||||| |||:||||| |||:|||||  
1283 GCACAGGCGCTAACAGA 1298

Search completed: May 26, 2003, 23:03:15  
Job time : 320 secs



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 22:54:23 ; Search time 58 seconds  
(without alignments)  
1903.511 Million cell updates/sec

Title: US-10-023-586B-2  
Perfect score: 1936  
Sequence: 1 MSLILLPSRSGSRGAL.....QNKLGHPAGRRKRCPLNRS 360

oring table: BLOSUM62 Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US10023586/runat\_22052003.114851.28697/app\_query.fasta\_1.519  
-DB=Issued\_Patents\_NA\_QFMT-fastbp\_SUFFIX=rni\_MINMATCH=0.1 -LOOPCL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10023586.ecgn\_1.1.40 -runat\_22052003.114851.28697 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	420.5	21.7	984	3	US-08-513-974B-41
2	420.5	21.7	1020	3	US-08-513-974B-370
3	420	21.7	1842	1	US-08-442-134A-1
4	420	21.7	1842	1	US-08-444-581B-1
5	420	21.7	1842	1	US-08-446-088A-1
6	413	21.3	984	3	US-08-513-974B-57
7	413	21.3	1023	3	US-08-513-974B-379
8	402	20.8	984	3	US-08-459-046-1
9	275	14.2	1301	2	US-08-467-948A-7
10	275	14.2	1301	3	US-08-467-947A-7
11	268	13.8	2160	4	US-08-188-275A-1
12	268	13.8	2162	4	US-09-351-198-1

13	268	13.8	2162	4	US-09-113-426-1	Sequence 1, Appl
14	267.5	13.8	1981	4	US-08-387-707-15	Sequence 15, Appl
15	267.5	13.8	1981	4	US-08-405-271A-15	Sequence 15, Appl
16	267	13.8	1610	3	US-08-889-108-7	Sequence 7, Appl
17	267	13.8	1610	5	PCT-US94-10358-7	Sequence 7, Appl
18	266.5	13.8	1901	1	US-08-153-848-43	Sequence 43, Appl
19	266.5	13.8	1901	3	US-09-299-843A-43	Sequence 43, Appl
20	266.5	13.8	1901	4	US-09-088-337B-43	Sequence 43, Appl
21	266.5	13.8	1901	5	PCT-US93-11153-43	Sequence 43, Appl
22	266.5	13.8	2453	5	PCT-US95-07180-1	Sequence 1, Appl
23	262.5	13.6	1205	1	US-08-417-103-13	Sequence 13, Appl
24	262.5	13.6	1634	1	US-07-816-283-1	Sequence 1, Appl
25	262.5	13.6	1634	1	US-08-417-103-1	Sequence 1, Appl
26	262.5	13.6	1910	2	US-09-009-438-1	Sequence 1, Appl
27	262.5	13.6	1910	3	US-09-207-493-1	Sequence 1, Appl
28	258.5	13.4	1640	3	US-08-781-250-1	Sequence 1, Appl
29	256	13.2	1265	1	US-07-816-283-3	Sequence 3, Appl
30	256	13.2	1265	1	US-08-417-103-3	Sequence 3, Appl
31	253.5	13.1	1102	2	US-08-742-440A-5	Sequence 5, Appl
32	253.5	13.1	1224	2	US-08-742-440A-4	Sequence 4, Appl
33	252	13.0	1255	1	US-08-097-938-3	Sequence 3, Appl
34	252	13.0	1255	1	US-08-476-000-3	Sequence 3, Appl
35	252	13.0	1255	1	US-08-472-840-3	Sequence 3, Appl
36	252	13.0	1255	2	US-08-476-976-3	Sequence 3, Appl
37	252	13.0	1255	3	US-08-474-410-3	Sequence 3, Appl
38	252	13.0	1255	4	US-08-486-673B-3	Sequence 3, Appl
39	250.5	12.9	1475	1	US-08-097-938-1	Sequence 1, Appl
40	250.5	12.9	1475	1	US-08-476-000-1	Sequence 1, Appl
41	250.5	12.9	1475	1	US-08-472-840-1	Sequence 1, Appl
42	250.5	12.9	1475	2	US-08-476-976-1	Sequence 1, Appl
43	250.5	12.9	1475	3	US-08-474-410-1	Sequence 1, Appl
44	250.5	12.9	1475	4	US-08-486-673B-1	Sequence 1, Appl
45	250.5	12.9	2732	1	US-08-476-000-60	Sequence 60, Appl

ALIGNMENTS

RESULT 1  
US-08-513-974B-41  
; Sequence 41, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiko  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01599  
; FILING DATE: 10-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-093989  
; FILING DATE: 19-AUG-1995



CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM: disk  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/513,974B  
 FILING DATE: 14-SEP-1995  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/JP95/01599  
 FILING DATE: 10-AUG-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 7-093989  
 FILING DATE: 19-AUG-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 7-057186  
 FILING DATE: 16-MAR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 7-007177  
 FILING DATE: 20-JAN-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 6-326611  
 FILING DATE: 28-DEC-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 6-270017  
 FILING DATE: 02-NOV-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 6-236357  
 FILING DATE: 30-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 6-236356  
 FILING DATE: 30-SEP-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 6-189274  
 FILING DATE: 11-AUG-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 6-189273  
 FILING DATE: 11-AUG-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 6-189272  
 FILING DATE: 11-AUG-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Resnick, David S.  
 REGISTRATION NUMBER: 34,235  
 REFERENCE/DOCKET NUMBER: 45753  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440  
 INFORMATION FOR SEQ ID NO: 370:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1020 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 28..1011  
 US-08-513-974B-370

Alignment Scores:  
 Pred. No.: 6,1e-35  
 Score: 420.50  
 Percent Similarity: 45.05%  
 Best Local Similarity: 32.14%  
 Query Match: 21.72%  
 DB: 3  
 Gaps: 9  
 Length: 1020  
 Matches: 117  
 Conservative: 47  
 Mismatches: 125  
 Indels: 76

US-10-023-586B-2 (1-360) x US-08-513-974B-370 (1-1020)  
 QY 29 MetGluLysValAspMetAsnThrSerGlnGluGlnGlyLeu-----CysGln 44  
 Db 28 ATGGAGCAG---GACAATGGCACCATCCAGGCTCCAGGCTGGCGGCCACACACCTGGCTG 84  
 QY 45 PheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIlePheIleLeu 64  
 Db 85 TACCGTGAGGATTCAAGCGACTGCTGCTAACCCCGGTATACTGGGTGGTGGTGGTGC 144  
 QY 65 GlyLeuProLeuAsnGlyThrValLeuThrPheHisSerTrpGlyGlnThrLysArgTrpSer 84  
 Db 145 GGCCTGCCACTGAACATCTGCGTCATCTCCGACATCTCCGCGACCTCCGCGACCTGACC 204  
 QY 85 CysAlaThrThrTyrLeuValAsnMetValAlaAspLeuLeuTyrValLeu---Leu 103  
 Db 205 CGTTCCGCTGTGTACACCTGAACCTGGCACTGGCGGACCTGATGATGCTGTTCACCTA 264  
 QY 104 ProPheLeuIleThrTyrSerLeuAspArgTrpPropheGlyGluLeuLeuCys 123  
 Db 265 CCCCTACTTATCTATACTAGCCAGAGGGGACCACTGGCCCTTCGGAGACCTGGCCTGC 324  
 QY 124 LysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCys 143  
 Db 325 CGCTTTGTAGGCTTCTCTTCTATGCAATCTACATGGCAGCATCTGTTCTCCACCTGC 384  
 QY 144 IleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg--- 162  
 Db 385 ATTAGCTTCAGCGCTACCTGGGCACTGCCACCCCTGGCTTCTCTGGCACAACCGTGA 444  
 QY 163 ThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuLeu 182  
 Db 445 GGTGCGCGTGTCTGTGGTAGTGTGGAGTCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 504  
 QY 183 LeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAsp 202  
 Db 505 CTGCCACGCGCAGTCTTTGTGTCACAGCATCCAGGCGCAACCCACTGTGTGTGTACGAC 564  
 QY 203 MetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSer 222  
 Db 565 CTGAGCCCAACCCATCTCTACTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 624  
 QY 223 GlyPheLeu---SerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluPro 242  
 Db 625 GGCTTCTTGTGCTGCCCTTCATAGCCTTACTGGCTTGTATTGCTG------ 668  
 QY 242 spGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHisPro 262  
 Db 669 -----CATGGCCCGCCGCTGTGTGCGCCAG 693  
 QY 262 spHisProThr----- 265  
 Db 694 GATGCCCAAGCAGGCTCTGTGGCCCAAGAGCGGCGCAGAGCGGCTCGTATGGCTGTG 753  
 QY 266 -----GlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeuLeu 284  
 Db 754 GTGGTGGCAGCTGTCTTGGCCATCAGCTT---CTGCTCTTCCACATCACCACAGACAGCC 810  
 QY 284 euProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCysG 304  
 Db 811 TACTTGGCTGTGCGCTCCAGCCCGGTGCTCTTGGCTGTGTGGTGGAGACCTTGGCTGCT 870  
 QY 304 LyLeuGlnAspMetGluAlaSerGlyGluCysGluGln----- 316  
 Db 871 GCCTACAAAGGCACTCGGCTTTCGCCAGTGTCAACAGTGTTCGGACCCCATCTCTCTTC 930  
 QY 317 -----LeuProGlnProSerProValLeuSerPheLysGlyGlyLys 331  
 Db 931 TACTTCACACACAGAGTTCGCGCGC-ANCCCCAGCATCTCTTACAGAGGCTCACAG- 988  
 QY 331 snArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisProAlaGly 351  
 Db 989 -----CCAAGTGGCA 998

QY 351 rgLysArg 353  
Db 999 GAGGCAGA 1006

RESULT 3

US-08-442-134A-1  
; Sequence 1, Application US/08442134A

; Patent No. 5596088

; GENERAL INFORMATION:

; APPLICANT: Boucher, Richard C.

; APPLICANT: Weisman, Gary A.

; APPLICANT: Turner, John T.

; APPLICANT: Harden, Thomas K.

; APPLICANT: Parr, Claude E.

; APPLICANT: Sullivan, Daniel M.

; APPLICANT: Erb, Laura

; APPLICANT: Lustig, Kevin D.

; TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bell, Seltzer, Park & Gibson

; CITY: Charlotte

; STATE: No. 5596088th Carolina

; COUNTRY: USA

; ZIP: 28234

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/442,134A

; FILING DATE: 16-MAY-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Sibley, Kenneth D.

; REGISTRATION NUMBER: 31,665

; REFERENCE/DOCKET NUMBER: 5470-71A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-420-2200

; TELEFAX: 919-881-3175

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1842 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 57..1181

; US-08-442-134A-1

Alignment Scores:

Pred. No.: 1.76e-34 Length: 1842

Score: 420.00 Matches: 127

Percent Similarity: 50.15% Conservative: 38

Best Local Similarity: 38.60% Mismatches: 140

Query Match: 21.69% Indels: 25

DB: 1 Gaps: 11

US-10-023-586B-2 (1-360) x US-08-442-134A-1 (1-1842)

QY 43 CysGlnPheSerGluLysTyrGlnValTyrLeuSerLeuAlaTyrSerIlellePhe 62

Db 129 TGCCGCTTCAACGAGGACTTCAAGTACGCTGCTGCTCTACGCGCTGGTGTGC 188

QY 63 IleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArg 82

Db 189 GTGCTGGGCTGTGTGACCCGCTGGGCTCTACATCTTCTTGTGGCGCTCAAGACC 248

QY 83 TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu 102  
Db 249 TGAATGCGTCCACCATATATGTTCCACCTGGCTGTCTGTCTGATGCACCTATGCGGCC 308  
QY 103 ---LeuProPheLeuIlelleThrTyrSerLeuAspArgTrpProPheGlyGluLeu 121  
Db 309 TCCCTGCGCGCTGCTGTATTACTACGCGCGCGGCGGCGCTTCCAGCAGGCGG 368  
QY 122 LeuCysLysValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeu 141  
Db 369 CTCTGCAAGCTGTGCGGCTTCTCTACACCAACCTTTACTGACAGCATCTCTTCCTC 428  
QY 142 ThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 161  
Db 429 ACCTGCATCAGCGTGCACCGGTGCTGGGCGCTTACGACCTCTCGCGTCCCTCGCTG 488  
QY 162 ArgThrArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGln 181  
Db 489 GCGCGGCGCGCTACGCTGCGCGGCTGGCGGCGCGCTGTGGTGTGTGGTGTGGCTG 548  
QY 182 LeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 201  
Db 549 CAGGCGCGCGCTGCTTCTTGTTCACCAACGCG---GCGCGCGCGCGCTACCTGCG 505  
QY 202 AspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyr---GlyIleValLeu 220  
Db 606 GACACCTGCGGCGCGGCTCTTACGCGCTTCTGCGCTACAGCTCAGTGTGCTGGCG 665  
QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
Db 666 TGTCTTCTGGCGGT---GCGCTTGGCGCTCATCTTGTCTGTACGTGCTCATGGCTG 724  
QY 241 ProAspGlnAlaArg-----GlyGluProHisGluAspArgGlnHisSerPro 256  
Db 725 ACTGCTAAAGCGGCTACGCGGCGCGCT-----GCTAGGCGGCGCA 772  
QY 257 SerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCys 276  
Db 773 GCGCAAGTCCGTCGCGCACCATCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 832  
QY 277 LeuProTyrHisSerLeuLeuLeuProHisHisLeuLeu-----SerAlaPhe 293  
Db 833 ATTCCACGCTCAC-----CCGACACCTCTACTACTCTCTCTCTCTCTCTCTCTCT 880  
QY 294 GlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGly 313  
Db 881 CAGCTGCCACACCTCAAGCCCATCAACATGGCTTACAGGTTTACCGGCT---GGCC 937  
QY 314 CysGluGlnLeuProGlnPro-SerProValLeuSerPheLysGlyLysAsnArgVa 333  
Db 938 TGCTAACAGTTG-----CCTTGACCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 991  
QY 333 lArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisProAla---GlyArg 352  
Db 992 ACGCTTGGCGCGAGATGCCAGCCA---CCCACTGGCGCGCGCGCTGCCACCGCGCTG 1048  
QY 352 sArgCysProGlyLeuAsnArgSer 360  
Db 1049 CCGCACGCTGGCGCTGCGCAGATCC 1073

RESULT 4

US-08-444-581B-1

; Sequence 1, Application US/08444581B

; Patent No. 5607836

; GENERAL INFORMATION:

; APPLICANT: Boucher, Richard C.

; APPLICANT: Weisman, Gary A.

; APPLICANT: Turner, John T.

; APPLICANT: Harden, Thomas K.

; APPLICANT: Parr, Claude E.

; APPLICANT: Sullivan, Daniel M.

; APPLICANT: Erb, Laura



[illegible]

Qy	257	SerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAla	276
		.....	
Db	773	GCSCAAGTCGGTCGCACCATCCCGCTGGTGGCTGCTTCCTGCCTCTCGCTCGCC	832
Ov	277	LeuProTrpHisSerLeuLeuLeuProHisHisLeuLeu-----SerAlaPheSer	293

Db 833 ATTCCAGTCACTACCTCTTCCGCTCGCTGGACCT 880

[illegible]

QY 314 CysGluGlnLeuProGlnPro-SerProValLeuSerPheLysGlyGlyLysAsnArgVa 333

Qy 333 lArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisProAla--GlyArgLy 352

Db 992 ACGTTTGGCCCGAGATGCCAAGCA--CCCACTGGCCCCAGCCCCGCCACCCCGGCTCG 104

Db 1049 CCGCACGCTGGGCCCTGGCAGATCC 1073

RESULT 6  
US-08-513-974B-57

Patent No. 6114139  
GENERAL INFORMATION:

; APPLICANT: Hosoya, Masaki  
 ; APPLICANT: Fujii, Ryo  
 ; APPLICANT: Ohtaki, Tetsuya

: APPLICANT: Fukusumi, Shoji  
 : APPLICANT: Ohgi, Kazuhiro  
 : TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN

; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF  
 ;  
 ; NUMBER OF SEQUENCES: 380  
 ;  
 ; CORRESPONDENCE ADDRESS:  
 ;

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston

STATE: MA  
COUNTRY: USA  
ZIP: 02109

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: 88768/313/3942  
 FILING DATE: 14-SEP-1995  
 CLASSIFICATION: 536

```

;
;
; APPLICATION NUMBER: PCT/JP95/01599
; FILING DATE: 10-AUG-1995
;

```

APPLICATION NUMBER: JP 7-093989  
FILING DATE: 19-AUG-1995

APPLICATION NUMBER: JP 7-057186  
FILING DATE: 16-MAR-1995

; APPLICATION NUMBER: JP 7-007177  
 ; FILING DATE: 20-JAN-1995  
 ;

APPLICATION NUMBER: JP 6-326611  
FILING DATE: 28-DEC-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-270017  
FILING DATE: 02-NOV-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236357  
FILING DATE: 30-SEP-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236356  
FILING DATE: 30-SEP-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189274  
FILING DATE: 11-AUG-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1994  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 57:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 984 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-513-974B-57

Alignment Scores:  
Pred. No.: 3,55e-34 Length: 984  
Score: 413.00 Matches: 126  
Percent Similarity: 48.61% Conservative: 31  
Best Local Similarity: 39.01% Mismatches: 146  
Query Match: 21.33% Indels: 22  
DB: 3 Gaps: 8

US-10-023-586b-2 (1-360) x US-08-513-974B-57 (1-984)

QY 33 AspMetAsnThrSerGlnGlnGlyLeu-----CysGlnPheSerGluLys 48  
Db 10 GACAAATGGACAGCCAGGCTCTGGGCTTGCACCCACACCTGTGTACCGCGAGAAC 69  
QY 49 TyrLysGlnValTyrSerLeuAlaTyrSerIleIleLeuGlyLeuProLeu 68  
Db 70 TTCAAGCAACTGTGCTGCCACCTGTATTTCGGCGTGTGGCGCTGCCCTGCCGCTG 129  
QY 69 AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88  
Db 130 AACATCTGTGTATTACCCAGATCTGCACGCTCCCGCGGCGCTGACCCGCGCGCGTG 189  
QY 89 TyrLeuValAlaMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 107  
Db 190 TACACCTAAACCTTGTCTGGGTGACCTGTATATGCCCTGCCCTGCCCTGCCCTGCATC 249  
QY 108 IleThrTrpSerLeuAspArgTrpProPheGlyLeuLeuLeuLeuValHis 127  
Db 250 TACAACCTATGCCAAGGTGATCACTGGCCCTTTGGCGACCTTCGCTGCCGCTGGTCCG 309  
QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147  
Db 310 TTCCTCTCTATGCAACCTGCACGCGCAGCATCTCTTCCTCACCTGCATCAGCTTCCAG 369  
QY 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166  
Db 370 CGCTACCTGGGCATCTGCCACCCCGCTGGCGCCCTGGGCACAAACCTGGGCGCGCGGCT 429

QY 167 AlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeu 186  
Db 430 GCCTGGCTAGTGTGTAAACCGTGTGGCTGGCCGTGACACCCAGTGCCTGCCACAGCC 489  
QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTrpAspMetThrSerGln 206  
Db 490 ATCTTTCCTGCCACAGCATCCAGCGTAAACGCGACTGTCTGTATGACCTCAGCCGCGCT 549  
QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226  
Db 550 GCCCTGGCCACCCACTATATGCCCTATGGCATGGCTCTCACTGTCTGCGGCTTCTGCTG 609  
QY 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246  
Db 610 CCCTTTCCTGCCCTGGCTGGCTGCTACTGCTCTCTGCGCTG---CCGCTGCGCGCAG 666  
QY 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThr 266  
Db 667 GATGGCCC-----GGCAGAGCCCTGTGGCCCGAGGCGGCGTGCAGGCGCGCGCATG 720  
QY 266 lyValTTrp-----ProLeu-HisProLeuPheCysAlaLeuProTyrHisSer 281  
Db 721 GCCG-TGGTGGTGGCTGCTGCTTTGCCATCAGCTT---CCTGCCCTTTTCATCATCACA 776  
QY 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301  
Db 777 GACAGCCTACCTGGCAGTGGCTCGACGCGCGGCGTCCCTGCACCTGTATTGGAGGCTT 836  
QY 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluGlnGlnLeuProGlnProSer 321  
Db 837 TGCAGCGCGCTTACAAAGGCGACGCGCGCTTTTGCAGTGCACAGCGTGTGGACCCCAT 896  
QY 322 ProValLeuSerPheLysGlyGlyLysAsnArgValArg-----LeuLeuGlnLys 338  
Db 897 CCTCTTCTAC--TTACCCAGAGAGTTCGCGCGCGGACCATGAGCTCCTACAGAAA 954  
QY 339 Leu 339  
Db 955 CTC 957

RESULT 7  
US-08-513-974B-379  
; Sequence 379, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01599

FILING DATE: 10-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-093989  
FILING DATE: 19-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-057186  
FILING DATE: 16-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-007177  
FILING DATE: 20-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-326611  
FILING DATE: 28-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-270017  
FILING DATE: 02-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236357  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236356  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189274  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 379:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1023 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
MOLECULE TYPE: cDNA  
NAME/KEY: CDS  
LOCATION: 37..1020  
US-08-513-974B-379

Alignment Scores:  
Pred. No.: 3.78e-34 Length: 1023  
Score: 413.00 Matches: 126  
Percent Similarity: 48.61% Conservative: 31  
Percent Local Similarity: 39.01% Mismatches: 146  
Query Match: 21.33% Indels: 22  
DB: 3 Gaps: 8

US-10-023-586B-2 (1-360) x US-08-513-974B-379 (1-1023)

Qy 33 AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluLys 48  
Db 46 GACATGTCAGGAGGCTGTGGCTTGGCCACCCACCCACCCACCTGTGTCTACCGCGAGAAC 105  
Qy 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuLeuProLeu 68  
Db 106 TTCAGCAACTGCTGCTGCCACTGTGTATTCGGCGGTGTGGCGGCTGGCGCTGCCGCTG 165  
Qy 69 AsnGlyThrValLeuThrPheSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88  
Db 166 AACATCTGTGTCATTACCATCTGCACGTCTCCCGCGGCTCCACCGGACCGCCGTG 225  
Qy 89 TyrLeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 107

Db 226 TACACCCTAAACCTTGCTCTGCTGCTACCTGCTATATGCTGCTGCTGCTGCTGCTCATC 285  
Qy 108 IleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127  
Db 286 TACAACATATGCCAAGGTGATCCTGCGCTTTGGCGCTTGGCGCTGCTGCTGCTGCTGCTG 345  
Qy 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuLeuThrCysIleSerValHis 147  
Db 346 TTCCTCTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 405  
Qy 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166  
Db 406 CGCTACCTGGGATCTGCCACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 465  
Qy 167 AlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuLeuLeuLeuProThrLeu 186  
Db 466 GCCTGGCTAGTGTGTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 525  
Qy 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206  
Db 526 ATCTTCGCTGCCACAGGATCCAGCGTAAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 585  
Qy 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--S 226  
Db 586 GCGCTGGCGCACCCACTATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 645  
Qy 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246  
Db 646 CCCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702  
Qy 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266  
Db 703 GATGGCCC-----GGCAGAGCTGTGGCCAGGCGGCTGGCAAGCGCGCCGCGCATG 756  
Qy 266 lyValTrp-----ProLeu-HisProLeuPheCysAlaLeuProThrHisSer 281  
Db 757 GCCG-TGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 812  
Qy 282 LeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySer 301  
Db 813 GACAGCTTACCTGGCGAGTGGGCTCGACGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 872  
Qy 302 GlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSer 321  
Db 873 TGCAGCGGCTACAAAGGCGACGCGCGCTTTGCCAGTGCACAGCGTCTGGACCCCAT 932  
Qy 322 ProValLeuSerPheLysGlyGlyLysAsnArgValArg-----LeuLeuGlnLys 338  
Db 933 CCTCTTCTAC---TTCACCCAGAAAGTTCCCGCGCGGCGGCGGCGGCGGCGGCGGCGG 990  
Qy 339 Leu 339  
Db 991 CTC 993  
RESULT 8  
US-08-459-046-1  
Sequence 1, Application US/08459046  
Patent No. 6008039  
GENERAL INFORMATION:  
APPLICANT: Coleman, Roger  
APPLICANT: Au-Young, Janice  
APPLICANT: Stuart, Susan G.  
TITLE OF INVENTION: A No. 6008039el Human Purinergic P2U Receptor  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3330 Hillview Avenue  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:



MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,046  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Luther, Barbara J.  
REGISTRATION NUMBER: 33954  
REFERENCE/DOCKET NUMBER: PF-0038 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-852-0195  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 984 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
IMMEDIATE SOURCE:  
LIBRARY: Placenta  
CLONE: 179696  
US-08-459-046-1

Alignment Scores:  
Pred. No.: 5 13e-33 Length: 984  
Score: 402.00 Matches: 119  
Percent Similarity: 46.60% Conservative: 32  
Best Local Similarity: 36.73% Mismatches: 150  
Query Match: 20.76% Indels: 24  
DB: 8 Gaps: 8

US-10-023-586B-2 (1-360) x US-08-459-046-1 (1-984)

QY 33 AspMetAsnThrSerGlnGlnGlyLeu-----CysGlnPheSerGluLys 48  
DB 10 GACAATGGCACAGACAGGCTGTGGCTTGCACCCACCACTGTGTACCGGAGAAC 69  
QY 49 TyrGlnValTyrLeuSerLeuAlaTyrSerLeuLeuPheLeuGlyLeuProLeu 68  
DB 70 TTCAGCACTGTGCTCCACCTGTGTATGGCGGTGTGGCGCTGCCCTCCCGCT 129  
QY 69 AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88  
DB 130 AACATCTGTGTCATTACCCAGATCTGCACGTCGCGCGGGCCCTGACCCGCGCGGTG 189  
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeu 107  
DB 190 TACACCTTAACCTTGTCTGCTGACCTGCTATGCTATGCTGCTGCTGCTGCTGCTGCT 249  
QY 108 IleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127  
DB 250 TACAACCTGCTCCCAAGTGATCACTGGCCCTTTGGGACTTCGCTGCCGCTGGTCCGC 309  
QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147  
DB 310 TTCCTCTTCATGCAACCTGCACGGGAGGATCTCTTCCTCCTACCTGCATCAGTTCCAG 369  
QY 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis 166  
DB 370 CGTACCTGGGCACTGCCACCCGCTGCGCCCTGGCACCAACCTGGGGGCGCGCGGCT 429  
QY 167 AlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeu 186  
DB 430 GCGTGGCTAGTGTGTAGCGGTGTGGCTGGCGGTGACACCCAGTCCCTGCCACAGCC 489  
QY 187 AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGln 206  
DB 490 ATCTTCGCTGCCACAGCATCCAGCGTAACCGCACTGCTGTGTATGACCTCAGCCCGCT 549

QY 207 GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeu--s 226  
DB 550 GCCCTGGCCACCCACTATATGCCCTATGGATGGCTCTCACTGTCACTGGCTTCTGCTG 609  
QY 226 erLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArg 246  
DB 610 CCCTTGTCTGCCCTGTGGCTGTACTGTCTCTCTGGCTG---CCGCTGTGGCGCAG 666  
QY 246 lyGluProHisGluAspArgGlnHisSerProSerGlnValHisProAspHisProThrG 266  
DB 667 GATGGCCC-----GGCAGACCTGTGGCCAGAGGCGGTGGCAGGCGCCCGCATG 720  
QY 266 ly-----ValTrpProLeuHisProLeuPheCysAlaLeuProTyrH 280  
DB 721 GCCGTGTGTGGCTGTCTTTGGCATCAGCTTCTGCGCTT-----TCACATC 771  
QY 280 isSerLeuLeuLeuProHisHisLeuSerAlaPheSerGlyLeuProAlaLeuAspG 300  
DB 772 ACCAAGACAGCCTACCTGGCAGTGGCTGCGACGCGCGGCGTCCCTGCACGTATTGGAG 831  
QY 300 lySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnP 320  
DB 832 GCCTTTCACGGGCTACAAAGCAGCGGCGCTTGCAGTGCACAGCGCTGCTGGAC 891  
QY 320 roSerProValLeuSerPheLysGlyGlyLysAsnArgValArg-----LeuLeuG 337  
DB 892 CCCATCCTCTCTAC--TTCACCCAGAAAGTTCCGCGCGGCGACCATGAGCTCCTAC 949  
RESULT 9  
US-08-467-948A-7  
Sequence 7, Application US/08467948A  
Patent No. 5998164  
GENERAL INFORMATION:  
APPLICANT: LI, YI  
APPLICANT: CAO, LIANG  
APPLICANT: NI, JIAN  
APPLICANT: GENTZ, REINER  
APPLICANT: BULT, CAROL J.  
APPLICANT: SUTTON III, GRANGER G.  
APPLICANT: ROSEN, CRAIG A.  
TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,948A  
FILING DATE: 30-MAR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04079  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: STEFFE, ERIC K.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

## ; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1301 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: both  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 161..1192  
US-08-467-948A-7

## Alignment Scores:

Pred. No.: 1.92e-19 Length: 1301  
Score: 275.00 Matches: 88  
Percent Similarity: 47.72% Conservative: 48  
Best Local Similarity: 30.88% Mismatches: 106  
Query Match: 14.20% Indels: 44  
DB: 2 Gaps: 10

US-10-023-586B-2 (1-360) x US-08-467-948A-7 (1-1301)

QY 23 GluGlyAlaSer-----ArgAspMetGluLysValAspMetAsnThrSer 37  
DB 122 GAAGTGGCTTCCAAACTGAAATGACGCTGCTTTACGATGGTAAGCGTTTACAGCTCC 181  
DB 38 GlnGluGlnGlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAla 57  
DB 182 CAC-----TGCTTCTATAATGACTCTTTAAGTACACATTGATGGTGCATG 229  
QY 58 TyrSerIlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheSerTrp 77  
DB 230 TTCAGCATGGTGTGTGGTGTAAATCAATGTGTGCCATATACATATTTTCATC 289  
QY 78 GlyGlnThrLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAsp 97  
DB 290 TCGGCTCCTCAAGTCGGAATGAACACTACACTTACATGATTACCTGCAATGTCAGAC 349  
QY 98 LeuLeuValLeu-----LeuProPheLeuIleIleThrTyrSerLeuAspArgTrp 116  
DB 350 TTGCTTTTGTGTGTACTTTTACCTTCAGGATTTT---TACTTCAACACGGAATTGG 406  
QY 117 ProPheGlyGluLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGly 136  
DB 407 CCATTGGAGATTACTTTGTAAAGATTTCTGTGCTGCTTTTATACCAACATGTACGGA 466  
QY 137 SerIleLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeu 156  
DB 467 ACATTCCTGCTTACCTGATTAAGTAGTAGATGATGATTTCTGCAATGCTTACCCATT 526  
QY 157 CysSerLeuProTyrArgThrArgHisAlaTrpLeuGlyThrSerThrThrTrpAla 176  
DB 527 AAGTCAAGACTCTAAGAACCAAAAGAAATGCAAGATGTTTGCACCTGGCGTGGTTA 586  
QY 177 LeuValValLeuGlnLeuLeuProThr-----LeuAlaPheSerHisThrAspTyrIle 194  
DB 587 ACTGTGATCGGAGGAGTGCACCCCGCTTTGTTGTACGCTCTACCCACTCTCAGGGTAAC 646  
QY 195 AsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGluAsnPheAsp-----210  
DB 647 ATGCCTCAGAGCTGCTTT-----GAAATTTTCCAGAACCCACATCG 691  
QY 211 ArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuGlyHis 230  
DB 692 AAAACATATCTCTCAAGGATTTGTAAT-----TTCATCGAAATAGTGGGATTT 739  
QY 231 Phe-----GlyValLeuPheThrAspGlyGln 239  
DB 740 TTTATTCCTCTAATTTTAATGTAATCTTGTCTAGTAGTGTGCTTAAACACTTTAACCAA 799  
QY 240 GluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnVal 259  
DB 800 CTGTTACATTAACTAGAGCAAAATAAACAACAACTAAGCTTTTAAATGATGTTTTGTGTA 859

QY 260 HisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyr 279  
DB 860 CATTG-GATCAT-----ATTCTGTTCTGTTTGTTCCTTACAATAT 900  
QY 280 HisSerLeuLeuLeu 284  
DB 901 CAATCTTATTATTATA 915

## RESULT 10

US-08-467-947A-7  
; Sequence 7, Application US/08467947A  
; Patent No. 6090575  
; GENERAL INFORMATION:

; APPLICANT: LI, VI  
; APPLICANT: CAO, LIANG  
; APPLICANT: NI, JIAN  
; APPLICANT: GENTZ, REINER  
; APPLICANT: BULT, CAROL J.  
; APPLICANT: SUTTON III, GRANGER G.  
; APPLICANT: ROSEN, CRAIG A.  
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein  
; TITLE OF INVENTION: Coupled Receptor GPRI  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVE., NW, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005

## COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,947A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/04079

FILING DATE: 30-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: STEFFE, ERIC K.

REGISTRATION NUMBER: 36,688

REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1301 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

TOPOLOGY: both

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 161..1192

US-08-467-947A-7

## Alignment Scores:

Pred. No.: 1.92e-19 Length: 1301  
Score: 275.00 Matches: 88  
Percent Similarity: 47.72% Conservative: 106  
Best Local Similarity: 30.88% Mismatches: 44  
Query Match: 14.20% Indels: 44  
DB: 3 Gaps: 10

US-10-023-586B-2 (1-360) x US-08-467-947A-7 (1-1301)

QY 23 GluGlyAlaSer-----ArgAspMetGluLysValAspMetAsnThrSer 37

```
Db 122 GAAAGTCTTCCAAACTGAAATTTGGACGTGCCTTTACGATGTGTAAGCGTTAAACAGCTCC 181
Qy 38 GlnGlnGlnGlyLeuCysGlnPheSerGlyLysTyrLysGlnValTyrLeuSerLeuAla 57
Db 182 CAC-----TGCTCTATAATGACTCCTTTAAGTACACTTGTGTATGGGTGCGATG 229
Qy 58 TyrSerIlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrPheHisSerTrp 77
Db 230 TTCAGCATGTGTGTGGTGTAAATATCCAAATTTGTGTGCAATATCATTTTCATC 289
Qy 78 GlyClnThrLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAsp 97
Db 290 TGGCTCCTCAAGTCGGAATGAACTACAACTACATGATTAATCTGGCAATGTCTAGAC 349
Qy 98 LeuLeuTyrValLeu---LeuProPheIleIleThrTyrSerLeuAspAspArgTrp 116
Db 350 TTGCTTTTGTGTATCTTACCTTCAGGATTTT---TACTTCACAAACGGAATGG 406
Qy 117 ProPheGlyGlnLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGly 136
Db 407 CCATTTGGAGATTTACTTTGTAGATTTCTGTGATGCTGTTTATACCAACATGTACGA 466
Qy 137 SerIleLeuLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeu 156
Db 467 AGCATTTCTGTCTTAACCTGTATTAGTGTAGATCGATTTCTGGCAATTTGTACCCATTT 526
Qy 157 CysSerLeuProTyrArgTrpArgHisAlaTrpLeuGlyThrSerThrThrTrpAla 176
Db 527 AAGTCAAGACTCTAAGAACCAAAAGAAATGCAAGATTTGTTGCACCTGGGTGTGTTA 586
Qy 177 LeuValValLeuGlnLeuLeuProThr-----LeuAlaPheSerHisThrAspTyrIle 194
Db 587 ACTGTGATCGGAGGAAGTGCACCGCGCTTTTGTTCAGTCTACCCACTCTCAGGGTAAC 646
Qy 195 AsnGlyGlnMetIleTyrTrpAspMetThrSerGlnGluAsnPheAsp----- 210
Db 647 AATGCCTCAGAGCCTGCTTT-----GAAATTTTCCAGAGCCACATGG 691
Qy 211 ArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuGlyHis 230
Db 692 AAAACATATCTCTCAAGGATGTAAAT-----TTCATCGAAATAGTGGGATTT 739
Qy 231 Phe-----GlyValLeuPheThrAspGlyGln 239
Db 740 TTTATCTCTAATTTTAAATGTAACTTGTCTAGTATGTGTCTTAAATACTTTAACCAAA 799
Qy 240 GluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnVal 259
Db 800 CCGTTACATTAGTAGAAGCAAAATAACAAAACTAAGGTTTAAATAATGATTTTGTGA 859
Qy 260 HisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyr 279
Db 860 CATTT-GATCAT-----ATTCTGTTCTGTTTGTCTTACATAT 900
Qy 280 HisSerLeuLeuLeu 284
Db 901 CAATCTTATTTTATA 915
RESULT 11
US-08-188-275A-1
; Sequence 1, Application US/08188275A
; Patent No. 6258556
; GENERAL INFORMATION:
; APPLICANT: Wang, Jia-Bel
; APPLICANT: Uhl, George R.
; APPLICANT: Johnson, Peter S.
; APPLICANT: Persico, Antonio
; TITLE OF INVENTION: cDNA and Genomic Clones Encoding Human
; TITLE OF INVENTION: Mu Opiate Receptor and the Purified Gene Product
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
```

```
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,275A
; FILING DATE: 28-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1173-449P
; TELEPHONE: 703-241-1300
; TELEFAX: 703-241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2160 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..2160
; OTHER INFORMATION: /label= cDNA
; OTHER INFORMATION: /note= "cDNA encoding human mu opiate receptor"
US-08-188-275A-1
Alignment Scores:
Pred. No.: 2,34e-18 Length: 2160
Score: 268.00 Matches: 95
Percent Similarity: 46.71% Conservative: 47
Best Local Similarity: 31.25% Mismatches: 128
Query Match: 13.84% Indels: 35
DB: 4 Gaps: 12
US-10-023-586B-2 (1-360) x US-08-188-275A-1 (1-2160)
Qy 58 TyrSerIlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrPheHisSerTrp 77
Db 441 TACTCCATCGTGTGGTGGGCTCTTCGGAACCTCCTCGGCACTGATGTATGTGATGTC 500
Qy 78 GlyClnThrLysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAsp 97
Db 501 AGATACACCAAGATGAGACTGCCACCAACATCATACATTTTCAACCTTGGCTGGCAGAT 560
Qy 98 LeuLeuTyrVal---LeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTrp 116
Db 561 GCCTTAGCCACCATGCTCCCTTCCAGAGTGTGAATTAC---CTAATGGGAACATGG 617
Qy 117 ProPheGlyGlnLeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGly 136
Db 618 CCATTTGGAACCATCTTTGCAAGATAGTAGTCTCCATAGATTACTATATAGATGTTACCC 677
Qy 137 SerIleLeuLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeu 156
Db 678 AGCATATTCACCTCTGCACCATGAGTGTGTGATCGATCATCATGCTGCCACCTGTC 737
Qy 157 CysSerLeuProTyrArgTrpArgHisAlaTrpLeuGlyThrSerThrThrTrpAla 176
Db 738 AAGGCTTAGATTTCCGTACTCCCGAAATGCCAAATATCAATGCTCTGCAACTGGATC 797
Qy 177 LeuValValLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGly 196
Db 798 CTCTCTTACGCCATTTGGTCTTCTCTTAATGTTCATGCTACACAAATACAGGCAAGGT 857
```

<hr/>					
Percent Similarity:	46.71%	Conservative:	47		
Best Local Similarity:	31.25%	Mismatches:	128		
Query Match:	13.84%	Indels:	35		
DB:	4	Gaps:	12		
<hr/>					
US-10-023-586B-2 (1-360) x US-09-351-198-1 (1-2162)					
Qy	58	TyrSerIleIlePheLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrp	77		
Db	441	TTTTTCTGTGGCGTGCGGTGGGGCTTCCTCGAAACATCTCCGTGCATGATGATTGTC	500		
Qy	78	GlyGlnThrLysArgTrpSerCysAlaThrThrTyrrLeuValAsnLeuMetValAlaAsp	97		
Db	501	AGATACCAAGATGAAGACTGCCACCACATCTACATTTTCAACCTTGCTCTGGCAGAT	560		
Qy	98	LeuLeuTyrVal---LeuLeuProPheLeuIleIleThrTyrrSerLeuAspArgTrp	116		
Db	561	GCCTTAGCCACCAGTAGTACCTCGCCCTCCAGAGTGTGAATTAC--CTAATGGGAACATGG	617		
Qy	117	ProPheGlyCluLeuLeuCysLysLeuValHisPheLeuPheTyrrIleAsnLeuTyrrGly	136		
Db	618	CCATTTGGAACCATCTCTTGCAAGATGATGATCTCCATGATAGTACTATAACATGTTTCAAC	677		
Qy	137	SerIleLeuLeuLeuThrCystIleSerValHisGlnPheLeuGlyValCysHisProLeu	156		
Db	678	AGCATATTACCTCTCGCACCATGATGTTGATCGATACATATGCAGTCTGCCACCTGTCT	737		
Qy	157	CysSerLeuProTyrrArgThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAla	176		
Db	738	AAGGCTTATGATTTCGGTACTCCCGAAATGCCAAATATCAATGCTGCAACTGGATC	797		
Qy	177	LeuValValLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrrIleAsnGly	196		
Db	798	CTCTCTTCAGCCATTGGTCTCTCTTAATGTTTCATGGCTACAACAATAACAGCAAGGT	857		
Qy	197	GlnMet-----IleTrpTyrrAspMethrSerGlnGlu	207		
Db	858	TCCATAGATTGTACATAACATTCTCTCATCCAACCTGGTAC-----TGGGAA	905		
Qy	208	AsnPheAspArgLeuPheAlaTyrrGlyIleValLeuThrLeuSerGlyPheLeuSerLeu	227		
Db	906	AACCTCGTGAAGATC-----TGTGTTTTTCATCTTCGCGCTTCATTTATGCCAGTG	953		
Qy	228	LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGlu	247		
Db	954	CTCAT-CATTACCCTGTGATGAGTCTGATGATCTTGCGCCTCAAGAGTGT-----	1003		
Qy	248	ProHis-----GluAspArgGlnHisSerProSerGlnValHisProAspHis	263		
Db	1004	CCGATGCTCTCTGGCTCCAAAGAAAGACAGGAATCTTCCAGGATCACCGAGTGGT	1063		
Qy	264	ProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrrHisSerLeuLeu	283		
Db	1064	GCTGGTGGTGGCTGTCTTCATCGTCTG-----CTGGACTCCCAATCA---CAT	1111		
Qy	284	LeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGlySerGlnCys	303		
Db	1112	TTACGTGCATCATTAAGCCCTTGGTTACAAT-----CCAGAAACTAGTMTCCAGACTGT	1165		
Qy	304	GlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnProSerProVal	323		
Db	1166	TTCTTGGCATCTTCGATTGCTCTAGGTTACACAACAGCTGCCTCAACCCAGTCTCT---	1222		
Qy	324	LeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuArgGlnAsnLys	343		
Db	1223	TTATGATTTCTGGATGAAAACTTCAAACG---ATGCTTTCAGAGATTCTGTATCCCAAC	1279		
Qy	344	LeuGlyGluHis	347		
Db	1280	CTCTTCCACAT	1291		
<hr/>					
RESULT 13					
US-09-113-426-1					



	Percent Similarity:	42.28%	Conservative:	49
	Best Local Similarity:	29.00%	Mismatches:	151
	Query Match:	13.82%	Indels:	63
	DB:	4	Gaps:	14
US-10-023-586B-2 (1-360) x US-08-387-707-15 (1-1981)				
Qy	24	GlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGlnGly	-----	41
Db	334	GGCTCCTGGCTCAACTTGTCCACAGTTTGCGAACCAGTCGCACCCATGGGTCCTAAC		393
Qy	42	-----LeuCysGlnPheSer	-----	49
Db	394	CCGACGGCCCTTGGGGAAACAGCAGCCTGTGCCCTCAGACGGCAGCCCTTCCATGGTC		453
Qy	50	LysGlnValTyrrLeuSerLeuAlaTyrrSerllelePheIleLeuGlyLeuProLeuAsn	69	
Db	454	ACAGCCATCACCATGGCCCTCTATTCTATCTGTGTGTAGTAGGGCTCTTTGAAC		513
Qy	70	GlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyrr	89	
Db	514	TTCCTGGTCATGTATGATGTATAGATATACCAATAAGACACTGCCACAACATCTAC		573
Qy	90	LeuValAsnLeuMetValAlaAspLeuLeuTyrrVal	---LeuLeuProPheLeuIleile	108
Db	574	ATTTTCAACCTTGTCTGGCAGATGCTTACCACTAGCAGCTGCCCTTTTCAGAGTGT		633
Qy	109	ThrTyrrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHisPhe	128	
Db	634	AACATAC---CTGATGGGAACGTGGCCCTTTGGAAACATCTCTCAAGATCGTATCTCA	690	
Qy	129	LeuPheTyrrIleAsnLeuTyrrGlySerIleLeuLeuThrCysIleSerValHisGln	148	
Db	691	ATAGACTACTACAACATGTTCCACAGTATCTTACCCTCTGCACATGAGTAGACCGC	750	
Qy	149	PheLeuGlyValCysHisProLeuCysSerLeuProTyrrArgThrArgArgHisAlatr	168	
Db	751	TACATTCGGCTGCCACCGGTCGAAGCCCTGGATTTCCCTACCCCCGAAATGCCAAA	810	
Qy	169	LeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe	188	
Db	811	ATTGTAATGTCGCACTGGATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	870	
Qy	189	SerHisThrAspTyrrIleAsnGlyGlnMet	-----Ile	199
Db	871	GCAACCAACAAATACAGCGAGGGTCCATAGATTGCACTTCTCTCATCCCACA	930	
Qy	200	TripyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrrGlyIleValLeu	219	
Db	931	TGGTAC-----TGGGAGAACCCTGCTCAAAATCTGTCTTCATCTTCGCTTC	978	
Qy	220	ThrLeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAsp	-----	237
Db	979	ATCATCGGGC-----CTCAT-CATCATGTGTGTGTATGAGCTGATCATCTT	1025	
Qy	238	--GlyGlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerPro	256	
Db	1026	ACACCTCAAGAGTCTCGCATGCTGTGGGCTCCAAGAAAGGACAGAA-----CCT	1079	
Qy	257	SerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAla	276	
Db	1080	GCGCAG-----GATCACCGGATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1130	
Qy	277	LeuProTyrrHisSerLeuLeuLeuProHisHisLeuSerAlaPheSerGlyLeuPro	296	
Db	1131	CTGACCCCCCATCATCTATGTATCATCAAAGACACTGATCATCATGATTCAGAAACCC	1190	
Qy	297	AlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGln	316	
Db	1191	TTTTCCAGAC-----TGTTTCTCGCACTTCTGCAATGCTGCTGCTGCTGCTGCTG	1241	
Qy	317	LeuProGlnProSerProValLeuSerPheLysGlyGlyLys	-----	330

[illegible]

Qy 70 GlyThrValLeuTrpHisSerTrpGlyClnThrLysArgTrpSerCysAlaThrThrTyr 89  
Db |||||  
514 TTCCTGTCATGTGATGTTAGATATACCAAAATGAAGACTGCCACCAACATCTAC 573  
Qy 90 LeuValAsnLeuMetValAlaAspLeuLeuTyrVal---LeuLeuProPheLeuIleIle 108  
Db |||||  
574 ATTTTCAACCTTGCTCTGGCAGATGCGCTTAGCCACTAGCAGCTGCGCTTTTCAGAGTGT 633  
Qy 109 ThrTyrSerLeuAspAspArgTrpPheGlyGluLeuLeuCysLysLeuValHisPhe 128  
Db |||||  
634 AACTAC---CTGATGGGAACCTGCGCTTGGAAACATCTCTCGAAGATCGTCTCA 690  
Qy 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148  
Db |||||  
691 ATAGACTACTACACATGTTTCCACAGTATCTTCCCTCTGCACCATGAGTGTAGACGC 750  
Qy 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgHisAlaTrp 168  
Db |||||  
751 TACATTCGCGTCTGCCACCGCTCAAGCCCTGGATTTCCTGACCCCGGAAATGCCAAA 810  
Qy 169 LeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuProThrLeuAlaPhe 188  
Db |||||  
811 ATTGTCATGCTGCAACTGGATCTCTCTCTGCCATTGGCTGCGCGTAAATGTTCTCATG 870  
Qy 189 SerHisThrAspTyrIleAsnGlyClnMet-----Ile 199  
Db |||||  
871 GCAACCAACAAATACAGCAGGGTCCATAGATTGCACCCCTCACGTTCTCTCATCCACA 930  
Qy 200 TrpTyrAspMetThrSerGlnGluAspPheAspArgLeuPheAlaTyrGlyValLeu 219  
Db |||||  
931 TGGTAC-----TGGGAGAACCTGCTCAAAATCTGTCTTCACTTCGCTTCGCTTC 978  
Qy 220 ThrLeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAsp----- 237  
Db |||||  
979 ATCATGCGGGC-----CTCAT-CATCACTCTGTGTTTGGACTGATGATCTT 1025  
Qy 238 ---GlyGlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerPro 256  
Db |||||  
1026 ACAGCTCAAGAGTCTCGCATGCTGCTGGGCTCCAAAGAAAGGACAGGAA-----CCT 1079  
Qy 257 SerGlnValHisProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAla 276  
Db |||||  
1080 GCGCAG-----GATCACCCGGATGGTCTGCTGGTCTGCTGCTGTTATTATTGCTG 1130  
Qy 277 LeuProTyrHisSerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuPro 296  
Db |||||  
1131 CTGGACCCCATCCACATCTATGTCATCATCAAGCACTGATCAGATTCAGAAACCCAC 1190  
Qy 297 AlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGln 316  
Db |||||  
1191 TTTCAGAC-----TGTTCTGCGACTTCTGCACTTGCCTTGGGTTACACAAACAG 1241  
Qy 317 LeuProGlnProSerProValLeuSerPheLysGlyLys----- 330  
Db |||||  
1242 CTGCTGAACCCAGTTCT---TTATGCGTCTCTGATGAAACTTCAACCATGTTTGTAG 1298  
Qy 331 -----AsnArgValArgLeuLeuGlnLysLeuArgGln 341  
Db |||||  
1299 AGAGTTCTGTCATCCCACTTCTCCACAATCGAAGACAGAAACTCTGCTCGAATCCGTCA 1358  
Qy 341 nAsnLysLeuGlyGluHisProAla 349  
Db |||||  
1359 AAAC---ACTAGGGAACACCCCTCC 1380

Search completed: May 27, 2003, 00:32:37  
Job time : 76 secs





GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 22:57:58 ; Search time 201 Seconds  
(without alignments)  
2365.007 Million cell updates/sec

Title: US-10-023-586b-2

Perfect score: 1936

Sequence: 1 MSLILLPSRSGSRGRRGAL.....QNKLGHPAGKRCPCPLNRS 360

coring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-Q-/cgn2\_1/USFTO\_spool/US10023586/runat\_22052003\_114852\_28717/app\_query.fasta.1.519  
-DB-Published-Applications\_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10023586 -CGN\_1\_1\_138 -runat\_22052003\_114852\_28717  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAPP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOPT=6 -FGAPOPT=7  
-FGAPEXT=7 -YGAPOPT=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications\_NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	1792	92.6	1017	9 US-09-885-453-3
2	921	47.6	510	9 US-09-782-974C-13
3	445.5	23.0	1429	9 US-09-077-173A-1
4	370	19.1	551	9 US-09-791-932-32

Alignment Scores: 1.47e-193 Length: 1017  
Pred. No.: 1792.00 Matches: 338  
Score: 99.71%  
Percent Similarity: 99.71%  
Conservative: 0

ALIGNMENTS

RESULT 1

US-09-885-453-3  
; Sequence 3, Application US/09885453  
; Publication No. US20030088080A1  
; GENERAL INFORMATION:  
; APPLICANT: Communi, Didier  
; TITLE OF INVENTION: RECEPTOR GPCRxl0  
; FILE REFERENCE: 9409/2082  
; CURRENT APPLICATION NUMBER: US/09/885,453  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 09/885,453  
; PRIOR FILING DATE: 2001-06-21  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1017  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DNA sequence  
; LOCATION: (1)..(1017)  
; OTHER INFORMATION: GPCRxl0 DNA sequence  
US-09-885-453-3

Sequence 85, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 7, Appl  
Sequence 2, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 1, Appl  
Sequence 12, Appl  
Sequence 2, Appl  
Sequence 15, Appl  
Sequence 8, Appl  
Sequence 26, Appl  
Sequence 6, Appl  
Sequence 25, Appl  
Sequence 13, Appl  
Sequence 12, Appl  
Sequence 1, Appl  
Sequence 2, Appl  
Sequence 9, Appl  
Sequence 23, Appl  
Sequence 22, Appl  
Sequence 3, Appl  
Sequence 11, Appl  
Sequence 51, Appl  
Sequence 51, Appl  
Sequence 1, Appl  
Sequence 15, Appl  
Sequence 15, Appl  
Sequence 4, Appl  
Sequence 17, Appl  
Sequence 17, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 16, Appl  
Sequence 24, Appl

Best Local Similarity: 99.71% Mismatches: 1  
Query Match: 92.56% Indels: 1  
DB: 9 Gaps: 0

US-10-023-586B-2 (1-360) x US-09-885-453-3 (1-1017)

```
QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
DB 1 ATGCTGTCCATTTTGTCTTCCAGGGGAAGCAGAAAGCGGAGCGTGTGGAGCTCTG 60
QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
DB 61 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATACATCAGAGAACAA 120
QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
DB 121 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 180
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrPheHisSerTrpGlyGlnThr 80
DB 181 ATCTTTATCTAGGCTGCGCACATAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 240
QY 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100
DB 241 AAGCGCTGAGCTGTCACCACTATCTGGTGAACCTGATGGCCGACCTGCTTTAT 300
QY 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120
DB 301 GTGCTATTGCCCTTCTCATCATCACTACTCACTAGATGACAGGTGGCCCTTCGGGGAG 360
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
DB 361 CTGCTCTGCAAGCTGTGCACTCTCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTG 420
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
DB 421 CTGACCTGCTCTCTGTGCACCACTCTCTGTAGGTGTGTCACCACTGTGTGCTGCTG 480
QY 161 TyrArgThrArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeu 180
DB 481 TACCGACCCGAGGATGCTGTGGTGGCCAGCACCACTGGGCCCTGTGTGCTC 540
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200
DB 541 CAGCTGTCTGCCACACTGCGCTCTCTCCACAGGACTACATCAATGGCCAGATGATCTGG 600
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
DB 601 TATGACATGACCAAGCAAGATTTTGTATCGCTTTTGTGCTACGGCATAGTCTTGACA 660
QY 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
DB 661 TTGTCTGGCTTTT-TCCCTCTGCTGCTATTTTGTGTGCTATTCATGATGTCAGGAG 719
QY 241 ProAspGlnAlaArgGlyCluProHisGluAspArgGlnHisSerProSerGlnValHis 260
DB 720 CTGTATCAAGCCAGAGGAGAACTCATGAGGACAGCAACACAGCCGCGGAGGTCCTAT 779
QY 261 ProAspHisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHis 280
DB 780 CCGGACCATCTCTACTGGTGTGGCTCTTTCACCTCTCTGTTTGTGGCTTCCATATCAC 839
QY 281 SerLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
DB 840 TCGCTCTCTTCTACCTACCATCTGCTTTTCTGCTTCTCAGGACTGCCAGCTCTTGATGC 899
QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
DB 900 AGCCAGTGTGGCTTCAAGATATGAGGCTCTGTGTAGTGTGAGCAGCTGCTCAACCC 959
QY 321 SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeu 339
DB 960 AGTCTGTACTTCTTCAAGGGGGGCAAAANTAGAGTACAGGCTCTCTCCAGAACTG 1016
```

RESULT 2

```
US-09-782-974C-13
; Sequence 13, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
; APPLICANT: Vogel, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
; FILE REFERENCE: 41USPHRM311
; CURRENT APPLICATION NUMBER: US/09/782,974C
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165,838
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: 60/198,568
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: 60/166,071
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/166,678
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 60/173,396
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/184,129
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 60/185,421
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/185,554
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: 60/186,530
; PRIOR FILING DATE: 2000-03-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 13
; LENGTH: 510
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-782-974C-13
```

Alignment Scores:  
Pred. No.: 3,16e-95 Length: 510  
Score: 921.00 Matches: 170  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 47.57% Indels: 0  
DB: Gaps: 0

US-10-023-586B-2 (1-360) x US-09-782-974C-13 (1-510)

```
QY 83 TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu 102
DB 1 TGGAGCTGTGGCCACCACCTATCTGGTGAACCTGATGGTGGCCGACCTGCTTAATGTGCTA 60
QY 103 LeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeu 122
DB 61 TTGCCCCCTTCTCATCATCACTACTACATGACAGGTGGCCCTTCGGGGAGCTGCTC 120
QY 123 CysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThr 142
DB 121 TGCAAGCTGTGCACCTCTCTTCTATATCAACCTTTACGGCAGCATCTGCTGTGCTGAC 180
QY 143 CysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg 162
DB 181 TGCATCTGTGCACAGTCTCTAGGTGTGCCACCACTGTGTGCTGCTGCTACCGG 240
QY 163 ThrArgArgHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeu 182
DB 241 ACCCGCAGSCATGCTGTGGTGGGCACCAACACCTGGGCCCTGTGTGCTCTCCAGCTG 300
```

Qy	85	CysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---	Leu 103
		:::	:::::
Db	385	GCAAGGCCACCTACATGTTCCACCTGGCATTCAGACACCTTGATGTGCTGCCTG	444
Qy	104	ProPheLeullelleThrTyrSerLeuAspAspArgTrpProPheGlyGluLeuLeuCys	123
		:::	:::      :::
Db	445	CCCACCCTACTACTATTATGAGCCCCACACCACTGGCCCTTTGGCACTCAGATCTGC	504
Qy	124	LysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCys	143
Db	505	AAGTTCGTCCGCTTCTCTTCTATTGGAACTCTACTGCAGTGTCTTTCTCCACCTGC	564
Qy	144	IleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThr	163
		:::::       :::::	:::::       :::::

ESUIT 3  
D00 A1CAGCGTGCACCGCTACCC1GCGGATC1GCGGACCCAC11CGGGGCTAC1ACCG1GGGGGCGCG  
000 A1CAGCGTGCACCGCTACCC1GCGGATC1GCGGACCCAC11CGGGGCTAC1ACCG1GGGGGCGCG 002

		164	ArgArgHisAlaLeuValLeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeu	183
Oy				
Db		625	CCTGCCTCGCAGCCCTCTCTGCCTGGCAGTTTGGTTGTTCGTAGCCGGCTGCCCTG	684
Oy		184	ProThrLeuAlaPheSerHisThrAspTyrIleAsnGlycInMetIleThrTyrrAspMet	203
Db		695	CCCACCTCTTCTTTGTCACACCACGACCAACAAGGGACCAACCCTCTCTGCGCATGACACC	744
Oy		204	ThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGly	223
Db		745	ACTCGGCTGAAGAGTTTGACCA-CTATGTGCA-----CTTCAGCTCGCGGT	791
Oy		224	-----pheLeuSerLeuLeuGlyHISpeGlyValLeuPheThrAspGly	238
			:	
Db		792	CATGGGGCTGCTCTTGGCGTGCCTGCCCTGGTGCACCTCTTGTGTTGTCATTGACCTATGCG	851

Qy	239	GlnGluProaspGlnAlaArgGly-----GluproHisGluAsp	251
Db	852	TcGTGGCGCTGTATCAGCCCTTGCAGCGCTGCACAGTCGTCTTCGCGCT	902
Qy	252	ArgGlnHisSerProSerGlnValHisProAspHisProThrGlyValTrpProLeuHis	271
		:::	
Db	903	-----CCGCTCTCTCCGCCACCATTAGCTGTGGTGCTGACTCTCTTGC	944
Qy	272	ProLeupheCysAlaLeuProTyrrHisSerLeuLeuLeuLeuLeuLeuLeuLeuLeu	289
Db	945	TGTTGTGCTTGTGCTTTCACATCAC-----CCGACCATTATTACTACTCTGC	992
Qy	290	--SerAlaPheSerGlyLeuProAlaLeuAaspGlySerGlnCysGlyLeuGlnAspMet	308
		:::	

Qy 309 GluAlaSerGlyGlyCysGluInLeuProGlnProSerProValLeuSerPheLysGly 328  
 Db 1053 TCGGCCCGCTGGCCAGTCCACAGCTGCCGTGGATCCCTGCTCTACTTGCTACTGGGGA 111  
 RESULT 4  
 US-09-791-932-32/c  
 ; Sequence 32, Application US/09791932  
 ; Publication No. US20030003451A1

1 GENERAL INFORMATION:  
 2 APPLICANT: Vogeli, Gabriel  
 3 APPLICANT: Parodi, Luis A.  
 4 APPLICANT: Hiebsch, Ronald R.  
 5 APPLICANT: Lind, Peter  
 6 APPLICANT: Kaytes, Paul S.  
 7 APPLICANT: Ruff, Valerie  
 8 APPLICANT: Huff, Rita M.  
 9 APPLICANT: Wood, Linda S.  
 10 TITLE OF INVENTION: No. US20030003451alel G Protein-Coupled Receptors  
 11 FILE REFERENCE: 00325 US1  
 12 CURRENT APPLICATION NUMBER: US/09/791,932  
 13 CURRENT FILING DATE: 2001-02-23  
 14 PRIOR APPLICATION NUMBER: 60/184,305  
 15 PRIOR FILING DATE: 2000-02-23  
 16 PRIOR APPLICATION NUMBER: 60/184,304

;  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184,303  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184,397  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/184,247  
; PRIOR FILING DATE: 2000-02-23  
; PRIOR APPLICATION NUMBER: 60/188,880  
; PRIOR FILING DATE: 2000-03-13  
; PRIOR APPLICATION NUMBER: 60/217,369  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/217,370  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/218,492  
; PRIOR FILING DATE: 2000-07-20  
; PRIOR APPLICATION NUMBER: 60/186,810  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: 60/188,064  
; PRIOR FILING DATE: 2000-03-09  
; PRIOR APPLICATION NUMBER: 60/186,457  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: 60/213,861  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 60/194,344  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 60/218,337  
; PRIOR FILING DATE: 2000-07-14  
; NUMBER OF SEQ ID NOS: 184  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 32  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-791-932-32

Alignment Scores:  
Pred. No.: 1.13e-32 Length: 551  
Score: 370.00 Matches: 74  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 19.11% Indels: 0  
DB: 9 Gaps: 0

US-10-023-586B-2 (1-360) x US-09-791-932-32 (1-551)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
Db 222 ATGCTGTCCATTTTCTTCTCCAGGGAGCAGAGCGGAGCGCTCTGGAGCTCTG 163  
Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGlu 40  
Db 162 CTCCTGGAGGAGCCTCCCGGACATGGAGAAGTGGACATCAATACATCAGGAACA 103  
41 GlyLeuGlnPheSerGlnLysThrLysGlnValThrLeuSerLeuAlaThrSerIle 60  
Db 102 GGTCTCTGCCAGTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 43  
Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyr 74  
Db 42 ATCTTTATCTAGGCTGCCACTAATATGCACTGCTCTTCTG 1

RESULT 5

US-09-782-974C-85  
; Sequence 85, Application US/09782974C  
; Publication No. US20030082534A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogeli, Gabriel  
; APPLICANT: Lind, Peter  
; APPLICANT: Wood, Linda S.  
; APPLICANT: Parodi, Luis A.  
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor  
; FILE REFERENCE: 41USPHM311  
; CURRENT APPLICATION NUMBER: US/09/782,974C

;  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: 60/165,838  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR APPLICATION NUMBER: 09/714,449  
; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: 60/198,568  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 60/166,071  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: 60/166,678  
; PRIOR FILING DATE: 1999-11-19  
; PRIOR APPLICATION NUMBER: 60/173,396  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/184,129  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: 60/185,421  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/185,554  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/186,530  
; PRIOR FILING DATE: 2000-03-02  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 85  
; LENGTH: 1020  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-782-974C-85

Alignment Scores:  
Pred. No.: 3.21e-22 Length: 1020  
Score: 281.50 Matches: 79  
Percent Similarity: 49.00% Conservative: 43  
Best Local Similarity: 31.73% Mismatches: 110  
Query Match: 14.54% Indels: 20  
DB: 9 Gaps: 6

US-10-023-586B-2 (1-360) x US-09-782-974C-85 (1-1020)

Qy 50 LysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeuAsn 69  
Db 97 AAGATGCACCTACCTCCCTGTTATTTATGTCATTTCTCTCTGGGATTTCCAGGCAAT 156  
Qy 70 GlyThrValLeuTyrHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyr 89  
Db 157 GCAGTAGTGATATCCACTTACATTTTCAAAATGAGACCTTGAAGAGCAGCACCATCAT 216  
Qy 90 LeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuIle 108  
Db 217 ATGCTGAACCTGGCCTGCACAGATCTGTATCTGATCACCAGCCTCCCTTCCTGATTAC 276  
Qy 109 ThrTyrSerLeuAspArgTrpPheGlyGluLeuLeuCysLysLeuValHisPhe 128  
Db 277 TACTATGCCAGTGGCGAAACTGGATCTTTGGAGATTTTCATGTAAAGTTTATCCGCTTC 336  
Qy 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148  
Db 337 AGCTTCCATTTCAACCTGTATAGCAGCATCTCTCTCCACCTGTTTTCAGCATCTCCGC 396  
Qy 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgHisAlaTyr 168  
Db 397 TACTGTGTGATCATTCACCAATAGCTCTTTTCCATTTCAAAAACACTCGATGTGCAATT 456  
Qy 169 LeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuPro---ThrLeuAla 187  
Db 457 GTAGCCTGTGCTGTGGTGATCATTTCCATGCTAGCTGTGATGCTCCGATGACCTTCTTG 516  
Qy 188 PheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGlnGlu 207  
Db 517 ATCACATCAACCAACAGGACCAAC---AGATCAGCCTGTCTCGACCTCACCAGTTCGGAT 573  
Qy 208 AsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeu 227

```
Db 574 GAACCTCAATAGTAACTGATTTGACCTCAAGTAC-TTCTGCTCC 632
QY 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProaspGlnAlaArgGlyGlu 247
Db 633 CTGGGGTAGTAGTACACTTTGGCTA-----TACCACGATTAT 668
QY 248 ProHisGluAspArg-----GlnHisSerProSerGlnValHisProasp 262
Db 669 CCACACTTTGACCCAGTGGAGTCAAACTGACAGCT-GCCTTAAGCAGAAAGCAGCAGGC 727
QY 263 HisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeu 282
Db 728 TAACCATTC---TGCTACTCTCTGCAATTT-TACGTATGTTTTTACCCCTCCATATCTTG 783
283 LeuLeuProHisLeuLeuSerAla 291
Db 784 AGGTCATTGAGTCAATCTCAGCC 810
```

## RESULT 6

```
US-10-023-775B-1
; Sequence 1, Application US/10023775B
; Publication No. US2003002282A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Ltd. (EP/GB) only)
; APPLICANT: Pfizer Inc. (US, JP, EB except GB)
; APPLICANT: Fidock, Mark David
; TITLE OF INVENTION: No. US2003002282A1el Polypeptide
; FILE REFERENCE: PC10959AGPR
; CURRENT APPLICATION NUMBER: US/10/023,775B
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: GB 0030854.4
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: US 60/260,590
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: US 60/296,660
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: GB 0111031.1
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-023-775B-1
```

```
Alignment Scores:
Pred. No.: 6,95e-22 Length: 1014
Score: 278.50 Matches: 76
Percent Similarity: 49.17% Conservative: 42
Best Local Similarity: 31.67% Mismatches: 105
Query Match: 14.39% Indels: 20
DB: 9 Gaps: 6
```

US-10-023-586B-2 (1-360) x US-10-023-775B-1 (1-1014)

```
QY 50 LysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheLeuGlyLeuProLeuAsn 69
Db 94 AAGATGCACTACCTCCCTGTTATTATGCGCATATCTCTCGTGGGATTTCCAGGCAAT 153
QY 70 GlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyr 89
Db 154 GCAGTAGTGATATCCACTTACATTTTCAAAATGAGACCTTGGAGACGACCATCATCATT 213
QY 90 LeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 108
Db 214 ATGCTGAACCTGGCCTGCACAGATCTGCTGTATCTGACACGCTCCCTCTCCCTGATTTCAC 273
QY 109 ThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHisPhe 128
Db 274 TACTATCCCAAGTGGCGAAACTGGATCTTTGGAGATTTCATGTAAAGTTATCCCGCTTC 333
```

```
QY 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148
Db 334 AGCTTCCATTCAACCTGTATAGCAGCATCTCTCTCCTCACTGTTTTCACGATCTTCCGC 393
QY 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgHisAlaTrp 168
Db 394 TACTGTGTGATCATTCACCAATGAGCTGCTTTTCCATTCCAAAACCTCGATGTGCGATT 453
QY 169 LeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuPro---ThrLeuAla 187
Db 454 GTACCTGTGCTGTGGTGGATCATTTCACTGTAGCTGTCTCATTCGATGACCTTCTTG 513
QY 188 PheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGlu 207
Db 514 ATCACAATCAACCAACAGGACCAAC---AGATCAGCCTGTCTCGACCTCACCAGTTCGAT 570
QY 208 AsnPheAspArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeu 227
Db 571 GAACCTCAATACTATTAGTGGTACACCTAATTTTGACTGCAACTAC-TTCTGCTCC 629
QY 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProaspGlnAlaArgGlyGlu 247
Db 630 CTGGGTGATAGTACACTTTGCTA-----TACCACGATTAT 665
QY 248 ProHisGluAspArg-----GlnHisSerProSerGlnValHisProasp 262
Db 666 CCACACTCTGACCCATGAGCTGCAAACTGACAGCT-GCCTTAAGCAGAAAGCAGCAGGC 724
QY 263 HisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeu 282
Db 725 TAACCATTC---TGCTACTCTCTGCAATTT-TACGTATGTTTTTACCCCTCCATATCTTG 780
RESULT 7
US-10-023-144-1
; Sequence 1, Application US/10270144
; Publication No. US20030049790A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000750CON
; CURRENT APPLICATION NUMBER: US/10/270,144
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/205,196
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Human
US-10-023-144-1
Alignment Scores:
Pred. No.: 6,95e-22 Length: 1014
Score: 278.50 Matches: 76
Percent Similarity: 49.17% Conservative: 42
Best Local Similarity: 31.67% Mismatches: 105
Query Match: 14.39% Indels: 20
DB: 9 Gaps: 6
US-10-023-586B-2 (1-360) x US-10-023-144-1 (1-1014)
QY 50 LysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheLeuGlyLeuProLeuAsn 69
Db 94 AAGATGCACTACCTCCCTGTTATTATGCGCATATCTCTCGTGGGATTTCCAGGCAAT 153
QY 70 GlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyr 89
Db 154 GCAGTAGTGATATCCACTTACATTTTCAAAATGAGACCTTGGAGACGACCATCATCATT 213
QY 90 LeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 108
```



US-09-885-453-2

## Alignment Scores:

Pred. No.: 6,95e-22 Length: 1014  
Score: 278.50 Matches: 76  
Percent Similarity: 49.17% Conservative: 42  
Best Local Similarity: 31.67% Mismatches: 105  
Query Match: 14.39% Indels: 20  
DB: 9 Gaps: 6

US-10-023-586B-2 (1-360) x US-09-885-453-2 (1-1014)

QY 50 LysGlnValTyrLeuSerLeuAlaTyrSerIlePheLeuGlyLeuProLeuAsn 69  
Db 94 AAGATGACCTACCTCCCTGTTATTTATGGCATTTATCTCTCGGGATTTCCAGGCAAT 153  
QY 70 GlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyr 89  
Db 154 GCAGTAGTGATATCCACTTACATTTTCAAAATGAGCTTGGAGAGCAGCACCATCAT 213  
QY 90 LeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 108  
Db 214 ATGCTGAACCTGGCGCTGCACAGATCTGCTATCTGACAGCTTCCCTCTCTGATTCAC 273  
QY 109 ThrTyrSerLeuAspArgTrpProPheGlyGlnLeuLeuCysLysLeuValHisPhe 128  
Db 274 TACTATGCCAGTGGCGAAACTGGATCTTTGGAGATTTTCAAGTTTATCCGCTTC 333  
QY 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148  
Db 334 AGCTTCCATTTCAACCTGTATAGCAGATCTCTCTCTCAGCTTTCAGCATCTTCCGC 393  
QY 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgHisAlaTyr 168  
Db 394 TACTGTGTATCATTCACCAATGAGCTGCTTTCCATTCACAAACTCGATGTGCAGTT 453  
QY 169 LeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuPro---ThrLeuAla 187  
Db 454 GTAGCTGTGCTGTGGTGGATTCATTTCACTGTAGCTGTCAATCCGATCCGCTTCTTG 513  
QY 188 PheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGlu 207  
Db 514 ATCATCAACCAACAGGACCAAC---AGATCAGCCTGTCTCGACCTCAGCATTCGGAT 570  
QY 208 AsnPheArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeu 227  
Db 571 GAACCTCAATACTATTAGTGTACAACTGATTTTGACTGCAACTAC-TTTCGCTCC 629  
QY 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGlu 247  
Db 630 CTGGTGATAGTGACACTTTTGCTA-----TACCACGATTAT 665  
QY 248 ProHisGluAspArg-----GlnHisSerProSerGlnValHisProAsp 262  
Db 666 CCACACTCTGACCCATGAGCTGCAAACTGACAGCT-CCCTTAAGCAGAAAGCAGGAGGC 724  
QY 263 HisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeu 282  
Db 725 TAACCATTC---TGCTACTCTCTGCATTT-TACGTATGTTTTTACCCCTCCCATATCTTG 780

RESULT 10

US-09-943-798-3

; Sequence 3, Application US/09943798  
; Patent No. US20020065215A1  
; GENERAL INFORMATION:  
; APPLICANT: Glaxo Group Limited  
; TITLE OF INVENTION: Polypeptide  
; FILE REFERENCE: OG1021  
; CURRENT APPLICATION NUMBER: US/09/943,798  
; CURRENT FILING DATE: 2001-08-31  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3

; LENGTH: 1014

; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-943-798-3

## Alignment Scores:

Pred. No.: 6,95e-22 Length: 1014  
Score: 278.50 Matches: 76  
Percent Similarity: 49.17% Conservative: 42  
Best Local Similarity: 31.67% Mismatches: 105  
Query Match: 14.39% Indels: 20  
DB: 10 Gaps: 6

US-10-023-586B-2 (1-360) x US-09-943-798-3 (1-1014)

QY 50 LysGlnValTyrLeuSerLeuAlaTyrSerIlePheLeuGlyLeuProLeuAsn 69  
Db 94 AAGATGACCTACCTCCCTGTTATTTATGGCATTTATCTCTCGGGATTTCCAGGCAAT 153  
QY 70 GlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyr 89  
Db 154 GCAGTAGTGATATCCACTTACATTTTCAAAATGAGCTTGGAGAGCAGCACCATCAT 213  
QY 90 LeuValAsnLeuMetValAlaAspLeuTyrValLeu---LeuProPheLeuIle 108  
Db 214 ATGCTGAACCTGGCGCTGCACAGATCTGCTATCTGACAGCTTCCCTCTCTGATTCAC 273  
QY 109 ThrTyrSerLeuAspArgTrpProPheGlyGlnLeuLeuCysLysLeuValHisPhe 128  
Db 274 TACTATGCCAGTGGCGAAACTGGATCTTTGGAGATTTTCAAGTTTATCCGCTTC 333  
QY 129 LeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGln 148  
Db 334 AGCTTCCATTTCAACCTGTATAGCAGATCTCTCTCTCAGCTTTCAGCATCTTCCGC 393  
QY 149 PheLeuGlyValCysHisProLeuCysSerLeuProTyrArgThrArgHisAlaTyr 168  
Db 394 TACTGTGTATCATTCACCAATGAGCTGCTTTCCATTCACAAACTCGATGTGCAGTT 453  
QY 169 LeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuPro---ThrLeuAla 187  
Db 454 GTAGCTGTGCTGTGGTGGATTCATTTCACTGTAGCTGTCAATCCGATCCGCTTCTTG 513  
QY 188 PheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyrAspMetThrSerGlnGlu 207  
Db 514 ATCATCAACCAACAGGACCAAC---AGATCAGCCTGTCTCGACCTCAGCATTCGGAT 570  
QY 208 AsnPheArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeu 227  
Db 571 GAACCTCAATACTATTAGTGTACAACTGATTTTGACTGCAACTAC-TTTCGCTCC 629  
QY 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGlu 247  
Db 630 CTGGTGATAGTGACACTTTTGCTA-----TACCACGATTAT 665  
QY 248 ProHisGluAspArg-----GlnHisSerProSerGlnValHisProAsp 262  
Db 666 CCACACTCTGACCCATGAGCTGCAAACTGACAGCT-CCCTTAAGCAGAAAGCAGGAGGC 724  
QY 263 HisProThrGlyValTrpProLeuHisProLeuPheCysAlaLeuProTyrHisSerLeu 282  
Db 725 TAACCATTC---TGCTACTCTCTGCATTT-TACGTATGTTTTTACCCCTCCCATATCTTG 780

RESULT 11

US-10-270-144-3

; Sequence 3, Application US/10270144  
; Publication No. US20030049790A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: C1000750CON





Qy 208 AsnProAspArgLeuPheAlaThrGlyIleValLeuThrLeuSerGlyPheLeuSerLeu 227  
Db 1117 GAACCTCAATACTAATAGTGCTACAACTAATTTTGACTGCACTACTTTGCCCTCCCTT 1176  
Qy 228 LeuGlyHisPheGlyValLeuPheThrAspGlyGlnGluProAspGlnAlaArgGlyGlu 247  
Db 1177 GGTGATAGTACACTTTGCTTACACGAT-----TAT 1209  
Qy 248 ProHisGluAsp 251  
Db 1210 CCACACTCTGAC 1221

RESULT 13

US-10-024-494-7

Sequence 7, Application US/10024494

Publication No. US20030044898A1

GENERAL INFORMATION:

APPLICANT: LI, YI  
CAO, LIANG  
ROSEN, CRAIG A.  
TITLE OF INVENTION: Human GPR4 G-Protein Coupled Receptor and  
Nucleotides Encoding Same (As Amended)  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESS: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/024,494  
FILING DATE: 21-Dec-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/465,973  
FILING DATE: 06-JUN-1995  
APPLICATION NUMBER: PCT/US95/04079  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: STERNE, ERIC K.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.1140001/EKS/HCC

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1301 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: CDNA  
FEATURE:

NAME/KEY: CDS

LOCATION: 161..1192

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-10-024-494-7

Alignment Scores:

Pred. No.: 2.55e-21 Length: 1301  
Score: 275.00 Matches: 88  
Percent Similarity: 47.72% Conservative: 48  
Best Local Similarity: 30.88% Mismatches: 106  
Query Match: 14.20% Indels: 44  
DB: 9 Gaps: 10

US-10-023-586B-2 (1-360) x US-10-024-494-7 (1-1301)

Qy 23 GluGlyAlaSer-----ArgAspMetGluLysValAspMetAsnThrSer 37  
Db 122 GAAGTGTCTTCAAACTGAAATTTGGACGTGCTTACGATGGTAAAGCGTTAAACAGCTCC 181  
Qy 38 GlnGluGlnGlyLeuGlySerGlnPheSerGlnValThrLeuSerLeuAla 57  
Db 182 CAC-----TGCTTCTATAATGACTCCCTTAAAGTACACTTTCTATGGGTGCATG 229  
Qy 58 TyrSerIleIlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuThrPheHisThr 77  
Db 230 TTCAGCATGGTGTGTGCTTGGGTTAATATCACTTGTGTGCTATGCTATGCTATGCTATG 289  
Qy 78 GlyGlnThrLysArgTyrSerCysAlaThrThrThrThrThrThrThrThrThrThr 97  
Db 290 TCGTCTCTCAAGTCCGAAATGAACTACACTTACATGATGATTAACCTTGGCAATGTCACAC 349  
Qy 98 LeuLeuTyrValLeu---LeuProPheLeuIleThrThrThrThrThrThrThrThrThr 116  
Db 350 TTGCTTTTGTGTTTACTTTACCTTTCAGGATTTT---TACTTCAACACACGGAATGG 406  
Qy 117 ProPheGlyGluLeuLeuGlySerGlnValHisPheLeuPheTyrIleAsnLeuTyrGly 136  
Db 407 CCATTGGAGATTTACTTTGTAAGATTTCTGTGATGCTGTTTATATACCAACATGTACGGA 466  
Qy 137 SerIleLeuLeuLeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeu 156  
Db 467 AGCATTTGTTCTTAACCTGATTAGTAGTATGATGATGATGATGATGATGATGATGATGAT 526  
Qy 157 CysSerLeuProTyrArgThrArgHisAlaThrLeuThrLeuGlyThrSerThrThrThr 176  
Db 527 AAGTCAAGACTCTAAGAACCAACCAAGAAATGCAAGATTTGTGCACTGGCGTGGTTA 586  
Qy 177 LeuValValLeuGlnLeuLeuProThr-----LeuAlaPheSerHisThrAspTyrIle 194  
Db 587 ACTGTGATCGGAGGAAGTGCACCCCGCTTTTGTGTTCTAGTCTACCCACTCTCAGGGTAA 646  
Qy 195 AsnGlyGlnMetIleTyrTyrAspMetThrSerGlnGluAsnPheAsp----- 210  
Db 647 AATGCTTCAGAGCCTGCTTT-----GAAATTTCCAGAACCCACATGG 691  
Qy 211 ArgLeuPheAlaTyrGlyIleValLeuThrLeuSerGlyPheLeuSerLeuGlyHis 230  
Db 692 AAACATATCTCTCAAGGATTTGTAAT-----TTTCATCGAAATAGTGGGATTT 739  
Qy 231 Phe-----GlyValLeuPheThrAspGlyGln 239  
Db 740 TTTATTCCTCTAATTTTAAATGTAACCTGTTCTAGTATGGTCTAAACCTTTAACCCAA 799  
Qy 240 GluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnVal 259  
Db 800 CCTGTTACATTAAGTAGAGCAAAATAAACAATAAGTGTAAATAATGATTTTGTGA 859  
Qy 260 HisProAspHisProThrGlyValThrProLeuHisProLeuPheCysAlaLeuProTyr 279  
Db 860 CATTT-GATCAT-----ATTCTGTTTCGTTTGTGTTCTCTACAAAT 900  
Qy 280 HisSerLeuLeuLeu 284  
Db 901 CAATCTTATTTTATA 915  
RESULT 14  
US-09-761-962-12  
; Sequence 12, Application US/09761962  
; Patent No. US20020077285A1  
; GENERAL INFORMATION:  
; APPLICANT: Memorial Sloan-Kettering Cancer Center  
; TITLE OF INVENTION: Identification and Characterization of Multiple Splice  
; TITLE OF INVENTION: Variants of Mu-  
; TITLE OF INVENTION: Opioid Receptor (MOR-1) Gene  
; FILE REFERENCE: 830002-2000.1  
; CURRENT APPLICATION NUMBER: US/09/761.962







GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 22:53:18 ; Search time 1772 Seconds  
(without alignments)  
3290.279 Million cell updates/sec

Title: US-10-023-586B.2  
Perfect score: 1936  
Sequence: 1 MLSTLLPSRSGSRGAL.....QNKLGHPAGKRCPLGNRS 360

coring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-Q/cgn2\_1/USPFO\_Spool/US10023586/runat\_22052003\_114851\_28657/app\_query.fasta\_i.519  
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10023586 -ACGN\_1\_1\_1906 @runat\_22052003\_114851\_28657 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_hc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description
1	414.5	21.4	2542	11	AK017378	AK017378 Mus muscu
2	414.5	21.4	3001	11	AK005013	AK005013 Mus muscu
3	378	19.5	851	13	BI833118	BI833118 603090834
4	360.5	18.6	641	14	BQ396255	BQ396255 NISC.ng19
5	359.5	18.6	638	9	AL675845	AL675845 AL675845
6	357	18.4	606	17	AZ953874	AZ953874 2M0219L17
7	355.5	18.4	1063	14	BM918491	BM918491 AGENCOURT
8	341	17.6	884	9	AL525099	AL525099 AL525099
9	337.5	17.4	898	13	BI819396	BI819396 603034571
c 10	325	16.8	1101	17	CNS05325	AL318925 Tetraodon
11	318	16.4	934	13	BI768397	BI768397 603053844
12	317	16.4	623	14	BQ038875	BQ038875 pgn1c.pk0
13	314	16.2	936	9	AL520218	AL520218 AL520218
14	310	16.0	801	9	AL549441	AL549441 AL549441
15	309.5	16.0	507	12	BF938665	BF938665 fm79g12.y
16	302	15.6	1118	14	BM918898	BM918898 AGENCOURT
17	300.5	15.5	899	14	BQ959110	BQ959110 AGENCOURT
18	295.5	15.3	484	14	BQ551383	BQ551383 H4008H11-
19	290	15.0	955	9	AL521440	AL521440 AL521440
c 20	286.5	14.8	408	13	BI401676	BI401676 MI-P-CP0-
21	276.5	14.3	589	14	BQ396804	BQ396804 NISC_dg23
22	274.5	14.2	781	12	BG681003	BG681003 602628822
23	272	14.0	546	12	BG552112	BG552112 ddi7a03.y
24	271.5	14.0	784	13	BI820686	BI820686 603034472
25	270	13.9	766	9	AJ450835	AJ450835 AJ450835
26	269	13.9	808	9	AJ456719	AJ456719 AJ456719
27	267	13.8	749	9	AJ447002	AJ447002 AJ447002
28	266.5	13.8	1083	14	BM805382	BM805382 AGENCOURT
29	266	13.7	747	9	AJ446611	AJ446611 AJ446611
30	266	13.7	773	9	AJ451396	AJ451396 AJ451396
31	266	13.7	798	9	AJ451155	AJ451155 AJ451155
32	266	13.7	900	14	BQ724849	BQ724849 AGENCOURT
33	265.5	13.7	806	17	CNS03VJW	AL262517 Tetraodon
34	265	13.7	800	9	AJ452624	AJ452624 AJ452624
35	265	13.7	860	9	AJ446632	AJ446632 AJ446632
36	265	13.7	890	9	AJ456561	AJ456561 AJ456561
37	264	13.6	741	9	AJ453402	AJ453402 AJ453402
38	263	13.6	740	9	AJ450921	AJ450921 AJ450921
39	262.5	13.6	845	9	AJ456135	AJ456135 AJ456135
40	262.5	13.6	879	14	BQ933137	BQ933137 AGENCOURT
41	260	13.4	1136	14	BM906977	BM906977 AGENCOURT
42	259	13.4	724	9	AJ447835	AJ447835 AJ447835
43	259	13.4	1064	13	BI517798	BI517798 603042051
44	258.5	13.4	642	12	BG159797	BG159797 de88h09.y
45	258	13.3	746	9	AJ450259	AJ450259 AJ450259

ALIGNMENTS

RESULT: 1  
AK017378  
LOCUS

DEFINITION  
Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15;purinergic receptor P2Y, G-protein coupled 2, full insert sequence.

ACCESSION  
AK017378.1  
VERSION  
AK017378.1  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (strain:C57BL/6J) 6 days neonate head cDNA to mRNA, clone.lib:RIKEN full-length enriched mouse cDNA library  
Clone:5430432J15.

ORGANISM  
Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Muridae; Murinae; Mus.

AK017378 2542 bp mRNA linear HTC 19-JAN-2002  
Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15;purinergic receptor P2Y, G-protein coupled 2, full insert sequence.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE

1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3

REFERENCE  
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, K.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,  
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,  
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
Kuehl, P., Lewis, S., Matsuo, Y., Nikaudo, I., Pesole, G.,  
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,  
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,  
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,  
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,  
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,  
Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,  
Toyooka, K., Wang, K. H., Westz, C., Whittaker, C., Wilming, L.,  
Yoshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.  
and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851  
5 (bases 1 to 2542)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,  
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,  
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,  
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,  
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,  
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,  
Saito, H., Saito, R., Sakai, K., Sakai, K., Sano, H., Sasaki, D.,  
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,  
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
Hayashizaki, Y.  
Direct Submission

TITLE  
JOURNAL

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
KANAGAWA 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,

COMMENT

Fax: 81-45-503-9216)  
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5', GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. cDNA went  
through one round of normalization to Rot = 10.0 and subtraction to  
Rot = 100.0. Second strand cDNA was prepared with the primer  
adaptor of sequence [5',  
GAGAGAGAGATCTCGAGTTAATTAATCCCCCCCCCC 3']. cDNA was cleaved  
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after  
bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3',  
end: BamHI. Host: DH10B.

FEATURES  
source

Location/Qualifiers  
1..2542  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="PANTOM\_DB:5430432J15"  
/db\_xref="MGD:MGI:1897121"  
/db\_xref="taxon:10090"  
/clone="5430432J15"  
/tissue\_type="head"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="6 days neonate"  
1..2542  
/gene="P2ry2"  
484..1605  
/gene="P2ry2"  
/note="data source: MGD, source key: MGI:105107,  
evidence: ISS  
Purine receptor P2Y, G-protein coupled 2  
putative"  
/codon\_start=1  
/protein\_id="BAB30719.1"  
/db\_xref="GI:12856589"  
/db\_xref="MGD:MGI:105107"  
/translation="MAADLEPNSTINGTWGDELGYKCRNEDEKYLPLVSYGVVC  
VLGCLNVALTIFLCRLKTNASTTIFHLAVSDLSLAASPLPLVLYYARGDHPFVS  
VLCKLRFLELYNLCISILFTCSIVHRCGLVRLPSLRGRARYARRAAVWVVL  
VLACAPVLYFTSVTRITCHDTSARELFSEFVAYSSVMLGLFLAVPESVILVCY  
VLMARLLPAYGTGGLPRAKRSVRTIALVLAFCFLPFLHTRLYYSFRLDL  
SCHTLNINMAYKITRPLASNSCLDPVLYFLAGORLVREARDAKPPTEPTSPQARR  
KLGLHRPNRTVRKDLSSVSDSRPTSTPAGSETKDIRL"

gene  
CDS

BASE COUNT 517 a 737 c 705 g 582 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 1-le-33 Length: 2542  
Score: 414.50 Matches: 132  
Percent Similarity: 48.66% Conservative: 50  
Best Local Similarity: 35.29% Mismatches: 155  
Query Match: 21.41% Indels: 38  
DB: 11 Gaps: 9  
US-10-023-586B-2 (1-360) x AK017378 (1-2542)  
Qy 2 LeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeuLeu 21  
|||||  
Db 435 CTGACATCCTC-----ACCACTCAAGACAGCAGGTGATCAG 473  
Qy 22 LeuGluGly-AlaSerArgAspMetGluLysValAspMetAsnThrSer----- 37  
|||||  
Db 474 GTCCAGGGCAATGGCAGCAGACCTGGACCCCTGGATACCATCAGCGACCTGGGA 533  
Qy 38 ----GlnGluGlnGlyLeu---CysGlnPheSerGluLysTyrLysGlnValTyrLeuSe 55  
|||||  
Db 534 GGGGACGACCTGGGATACAGTGTCTTCAACGAGGACTTCAAGTACGTCTGTGGCC 593







QY	108	IleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis	127	
DB	356	TACAACATATGGCCCAAGGTGATCAGTGGCCCTTTGGCGACTTCGGCTGGCGCTGGTCCGC	415	
QY	128	PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuThrCysIleSerValHis	147	
DB	416	TTCCCTCTTGTATGCCAACCTGCAGCGCAGCATCCCTTCCTCACTGCATCAGCTTCAG	475	
QY	148	GlnPheLeuGlyValCysHisProLeuLeuTyrProTyrArg	166	
DB	476	CGCTACCTGGCATCGCACCCGCTGGCCCTGGCAACACGTGGGGCGCCGGCT	535	
QY	167	AlaTrpLeuGlyThrSerThrTrpAlaLeuValLeuGlnLeuLeuProThrLeu	186	
DB	536	GCTGGCTAGTGTGTAGCCGTGGTGGCCGTGACAAACCATGTCCTGCCACAGCC	595	
QY	187	AlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyrTyrAspMetThrSerGln	206	
DB	596	ATCTTCGCTGCCACAGGATCCAGCGTAACCGCACTGTCTGTATGACCTCAGCCCGCT	655	
QY	207	GluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuLeuSerGlyPheLeu	225	
DB	656	GCCCTGCCCAACCACCTATATGATGCCTATGCGATGGCTCTCACTGTCATCGCTCTCTG	712	
RESULT 4				
LOCUS	BQ396255	641 bp mRNA linear	EST 22-MAY-2002	
DEFINITION	NISC_ng19q07.y1_NICH	XGC Emb6 silurana tropicalis cDNA clone		
ACCESSION	BQ396255			
VERSION	BQ396255.1	GI:21083932		
KEYWORDS	EST.			
SOURCE	western clawed frog.			
ORGANISM	Silurana tropicalis			
REFERENCE				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE	Xenopodinae; Silurana.			
JOURNAL	1 (bases 1 to 641)			
COMMENT	NIH-XGC http://image.llnl.gov/image/html/xenopuslib_info.shtml. National Institute of Child Health and Human Development, National Cancer Institute, Xenopus Gene Collection Unpublished (2002) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov cDNA Library Preparation: DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC) Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov Plate: LLAM11977 row: N column: 13 Seq primer: M13Rpl reverse primer (ABI). Location/Qualifiers 1. .641 /organism="Silurana tropicalis" /db_xref="taxon:8364" /clone="IMAGE:5383884" /clone_lib="NICH XGC Emb6" /tissue_type="neurula" /dev_stage="embryo, stages 14-19" /lab_host="DHI0B (phage-resistant)" /note="Vector: pCMV-SPORT6.cdb; Site.1: NotI; Site.2: EcoRV; Cloned unidirectionally. Primer: Oligo dt. Average insert size 2.1 Kb. Constructed by invitorgen. Note: This is a Xenopus Gene Collection (XGC) library."			
BASE COUNT	132 a	184 c	153 g	172 t
ORIGIN				
Alignment Scores:				
Pred. No.:	1.03e-28	Length:	641	
Score:	360.50	Matches:	77	
Percent Similarity:	55.61%	Conservative:	32	



```

SOURCE
ORGANISM      house mouse.
               Mus musculus
REFERENCE
AUTHORS       Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
               Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
               M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
               and Wright,D., Weiss,R.
TITLE         Mouse whole genome scaffolding with paired end reads from 10kb
               plasmid inserts
JOURNAL
COMMENT       Unpublished (2000)
               Contact: Robert B. Weiss
               University of Utah
               Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
               84112, USA
               Tel: 801 585 5606
               Fax: 801 585 7177
               Email: ddunn@genetics.utah.edu
               Insert Length: 10000 Std Error: 0.00
               Plate: 0219 row: L column: 17
               Seq primer: CGTTGTAACGACGGCCAGT
               Class: plasmid ends
               High quality sequence stop: 606.
FEATURES
source
1..606
   /organism="Mus musculus"
   /strain="C57BL/6J"
   /db_xref="taxon:10090"
   /clone_lib="Mouse 10kb plasmid UUGC2M library"
   /sex="Female"
   /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
   /note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT    100 a 203 c 145 g 158 t
ORIGIN
Alignment Scores:
Pred. No.:      2 28e-28      Length:      606
Score:          357.00      Matches:      74
Percent Similarity: 57.84%      Conservative: 33
Best Local Similarity: 40.00%      Mismatches: 75
Query Match:    18.44%      Indels:      3
DB:             17          Gaps:      2

US-10-023-586B-2 (1-360) x A2953874 (1-606)

Qy 44 GlnPheSerGluLysTrpLysGlnValTrpLeuSerLeuAlaTyrSerIlePhe163
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 19 GAACATGATCCGATTAAAGCACTGCTAACCCTGCTATCGGTGGTGGTG 78
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy 64 LeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrp 83
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 79 GTCCGCTGCCACTGAACATCTGCGTCAITGCCATATCTGGGCATCCCGCGGACCTG 138
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

84 SerCysAlaThrThrTrpLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu--- 102
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 139 ACCGTTCCGCTGTGTACACCTGAACCTGGCACTGGCGGACCTGATGATGCCTGTTC 198
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy 103 LeuProPheLeuIleIleThrTrpSerLeuAspArgTrpProPheGlyGluLeuLeu 122
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 199 CTACCCCTACTATCTATATACTAGCCAGATGGGACCACTGGCCCTTCGGAGACCTGCC 258
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy 123 CysLysLeuValHisPheLeuPheTrpIleAsnLeuTyrGlySerIleLeuLeuLeuThr 142
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 259 TGGCGCTTTGTACGCTCTCTTATGCCAATCTACATGGCAGCATCTGTCTCTCACC 318
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy 143 CysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg 162
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 319 TGCATTAGCTTCCACGCGTACCTGGGCATCTGCCACCCCTGGCTTCTGGCACAGCGT 378
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy 163 ---ThrArgArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGln 181
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 379 GGAAGTCGCGCTGCTGCTGGGTAGTGTGGAGTCATGTGGCTGTGCAGAGCCCAT 438
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy 182 LeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyr 201
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 439 TGCCTGCCACGCGAGTCTTTGCTGCCACATGCATCCATCGCAACCGCACCTGTGGTAC 498
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy 202 AspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyr-GlyIleValLeuThrLe 221
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 499 GACCTGAGCCACCCATCTCTGTCTACTGCTACCTGCGCTATGATGATGGCCCTCACGGT 558
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Qy 221 uSerGlyPheLeu 225
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 559 CATCGGCTCTTTG 571
   ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 7
BM918491
LOCUS      BM918491
DEFINITION AGENCOURT_6708833 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:574743
5', mRNA sequence.
ACCESSION  BM918491
VERSION     BM918491.1 GI:19368870
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1063)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: The I.M.A.G.E. Consortium, Inc.
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LUAM12773 row: j column: 20
              High quality sequence stop: 656.
FEATURES
source
1..1063
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="IMAGE:574743"
   /clone_lib="NIH_MGC_120"
   /lab_host="DH10B"
   /note="Organ: pooled pancreas and spleen; Vector:
pcmv-SPORT6; Site_1: NOT1; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is

```

normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

BASE COUNT 172 a 379 c 288 g 223 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 7.8e-28 Length: 1063  
Score: 355.50 Matches: 105  
Percent Similarity: 43.00% Conservative: 27  
Best Local Similarity: 34.20% Mismatches: 134  
Query Match: 18.36% Indels: 41  
DB: 14 Gaps: 7

US-10-023-586B-2 (1-360) x BM918491 (1-1063)

QY 58 TyrSerIlePheLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrp 77  
Db 238 TATTCGGCGGTGGCGGCTGGCGCTGCGCGCTGAACATCTGTCTCAATACCCAGATCTGC 297  
QY 78 GlyGlnThrLysArgTrpSerCysAlaThrThrThrLeuValAsnLeuMetValAlaAsp 97  
Db 298 ACCTCCCGCGGCGCTGACCGCGCGCTGTACACACCTAAACCTTCTCTGGCTGAC 357  
QY 98 LeuLeuTrpValLeu---LeuProPheLeuIleThrThrSerLeuAspAspArgTrp 116  
Db 358 CTGCTATATGCTGCTGCTGCTGCTGCTGCTATCTACAACTATGCCAAGGTGATCACTGG 417  
QY 117 ProPheGlyGluLeuLeuCysLysLeuValHisPheLeuPheTrpIleAsnLeuTrpGly 136  
Db 418 CCCTTTGGGCACTTCGCTGCGCGCTGGTGGCTTCTTCTATGCCAACCTGCACGGC 477  
QY 137 SerIleLeuLeuThrCysLysSerValHisGlnPheLeuGlyValCysHisProLeu 156  
Db 478 AGCATCTCTCTCACTACCTGATCAGCTCCAGCGCTACTGGGCATCTGCCACCGCTG 537  
QY 157 CysSerLeuProTrpArg---ThrArgArgHisAlaTrpLeuGlyThrSerThrTrp 175  
Db 538 GCGCCCTGGCAACAGCTGGGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597  
QY 176 AlaLeuValValLeuGlnLeuLeuProThrLeuAlaPheSerHisThrAspTrpIleAsn 195  
Db 598 CTGGCGCTGACACACCGAGTGCCTGCCACGACCATCTTCTGCTGCCACAGCATCCAGCGT 657  
QY 196 GlyGlnMetIleTrpTrpAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTrp 215  
Db 658 AACCGCACTGTCTGTATGATGACCTGACCGCGCTGCGCTGCGCTGCTGCTGCTGCTAT 717  
QY 216 GlyIleValLeuThrLeuSerGlyPheLeu---SerLeuLeuGlyHisPheGlyValLeuP 235  
Db 718 GGCATGGCTCTCACTGCTATCGGGCTCTGCTGCGCTTGTGCTGCTGCTGCTGCTGCTGCT 777  
QY 235 heThrAspGlyGlnGluProAspGlnAlaArgGlyGluProHisGluAspArgGlnHis 255  
Db 778 TGTCTCTGGCTGCGCTGTGGCGCGCGAGATGGCGCGCGAGAA-----824  
QY 255 exProSerGlnValHisPro-----AspHisProThrGlyValTrpProLeuH 271  
Db 825 -----CCTGTGGCGCGCAAGCGCGCTGGCGAGCGCGCGCGCGCGCATGNCCTGCT 876  
QY 271 isProLeuPheCysAlaLeuProTrpThrHisSerLeuLeuLeuProHisHisLeuLeuSerA 291  
Db 877 GTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 936  
QY 291 laPheSerGlyLeuProAlaLeuAspGlySerGlnCysGlyLeuGlnAspMetGluAla 311  
Db 937 CCTACCTGGCAATGGCGCTGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 993  
QY 311 erGlyGluCysGlnLeuPro-----GlnProSerProValLeuSerPheLysGlyC 329  
Db 994 AAGGGCGCTTGGCACCGCGCGCTTACAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1040  
QY 329 lYlYsAsnArgValArgLeuLeuGlnLysLeuArgGlnAsnLysLeuGlyGluHisProA 349

Db 1041 -----  
QY 349 laGlyArgLysArgCys 354  
Db 1045 TTGGCCCAAAAGGTGC 1061  
RESULT 8  
AL525099  
LOCUS AL525099 LTI\_NFL003\_NBC3 884 bp mRNA linear EST 13-FEB-2001  
DEFINITION AL525099 LTI\_NFL003\_NBC3 Homo sapiens cDNA clone CS0DC003YM02 5  
prime, mRNA sequence.  
ACCESSION AL525099  
VERSION AL525099  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 884)  
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
FEATURES  
Location/Qualifiers  
source 1..884  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DC003YM02"  
/clone\_lib="LTI\_NFL003\_NBC3"  
/sex="male"  
/tissue\_type="neuroblastoma cells"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
vector. Library was normalized. Library was constructed  
by Life Technologies. Contact : peng liang life  
Technologies, a division of invitrogen 9800 Medical Center  
Drive Rockville, Maryland 20850, USA fax : (1) 301 610  
8371 Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"

BASE COUNT 153 a 302 c 250 g 179 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.24e-26 Length: 884  
Score: 341.00 Matches: 74  
Percent Similarity: 55.15% Conservative: 17  
Best Local Similarity: 44.85% Mismatches: 68  
Query Match: 17.61% Indels: 6  
DB: 9 Gaps: 3

US-10-023-586B-2 (1-360) x AL525099 (1-884)

QY 33 AspMetAsnThrSerGlnGluGlyLeu-----CysGlnPheSerGluLys 48  
Db 389 GACAATGGCACAGCGCAGGCTCTGGCTTGGCCACCCACCATCTGCTTACCGCGAGAAC 448  
QY 49 TyrLysGlnValTrpLeuSerLeuAlaTrpSerIlePheIleLeuGlyLeuProLeu 68  
Db 449 TTCAGCACTCTCTGCTGCGCACCTGTGTATTGCGGGCTGCTGGCGCTGCGCTGCGCTG 508  
QY 69 AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThr 88  
Db 509 AACATCTGTGTATACCATCTGCAGCTGCCGCGCGCGCTGACCCGCGCGCGCTG 568  
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTrpValLeu---LeuProPheLeu 107

```

Db 569 TACACCTAAACCTTGCTCTGGCTGACCTGCTATATGCTCCCTGCCCTGCTCATC 628
Oy 108 lIeThrTySerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127
Db 629 TACAACTATGCCCAAGGTGATCACTGGCCCTTTGGCGACTGCGCTGCCGCTGGTCCGC 688
Oy 128 PheLeuPheTyIleAsnLeuTyGlySerIleLeuLeuLeuThrCysIleSerValHis 147
Db 689 TTCCTCTTCTATGCAACCTGCACGGCAGCATCTCTCTCTCCCTGACATCAGCTTCCAG 748
Oy 148 GlnPheLeuValCysHisProLeuLeuCysSerLeuProTyArg---ThrArgArghis 166
Db 749 CGTACCTGGGCACTGCACCGCTGGCCCTGGCGGCAACAGCTGGGGCGCGGCT 808
Oy 167 AlaTrpLeuGlyThrSerThrThrAlaLeuValLeuGlnLeuLeuProThrLeu 186
Db 809 GCCTGGCTAGTGTGTAGCTGTGGCTGGCGGTGACAAACCCAGTGGCTGCCACAGCC 868

Oy 187 AlaPheSerHisThr 191
Db 869 ATCTCGCTGCCACA 883

RESULT 9
BI819396
LOCUS      898 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603034571f1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175761 5',
            mRNA sequence.
ACCESSION  BI819396
VERSION     BI819396.1 GI:15930946
KEYWORDS   EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 898)
REFERENCE   NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LNAW11437 row: n column: 18
            High quality sequence stop: 884.
            Location/Qualifiers
                1..898
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5175761"
                /clone_lib="NIH_MGC_115"
                /lab_host="DH10B"
                /note="Organ: pooled brain, lung, testis; Vector:
                PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                source anonymous pool of 6 male brains, age range 23-27; 1
                male lung, age 27; and 1 male testis, age 69. Library is
                oligo-dr primed and directionally cloned (EcoRV site is
                destroyed upon cloning). Average insert size 1.8 kb,
                insert size range 1-3 kb. Library is normalized and
                enriched for full-length clones and was constructed by C.
                Gruber (Invitrogen). Research Genetics tracking code
                021. Note: this is a NIH_MGC Library."
BASE COUNT  149 a 303 c 255 g 191 t
ORIGIN

Alignment Scores: 5.51e-26 Length: 898
Pred. No.: 337.50 Matches: 83
Score: 51.96%
Percent Similarity: Conservative: 23

```

```

Best Local Similarity: 40.69% Mismatches: 87
Query Match: 17.43% Indels: 12
Db: 13 Gaps: 4
US-10-023-586B-2 (1-360) x BI819396 (1-898)

Oy 33 AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluLys 48
Db 201 GACAATGGCACAGGCGAGCTCTGGCTTGGCCACCCACCACTGCTGTCTACCGCGAGAAC 260
Oy 49 TyrLysGlnValTySerLeuSerLeuAlaTySerIleIlePheIleLeuGlyLeuProLeu 68
Db 261 TTCAGCAACCTGCTGCTGCCACCTGTGTATTGGGGGTGCTGGCGGCTGGCTGCCGGTG 320
Oy 69 AsnGlyThrValLeuThrPheHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88
Db 321 AACATCTGTGTATACCCAGATGTGCAGCTCCCGCGGGCGCTGACCGCGAGCGCGCTG 380
Oy 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyValLeu---LeuProPheLeuIle 107
Db 381 TACACCTTAAACCTTGCTCTGGCTGACCTGCTATATGCTCTCCCTGCCCTGCTCATC 440
Oy 108 lIeThrTySerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127
Db 441 TACAACATATGCCCAAGGTGATCACTGGCCCTTTGGCGACTTGGCGCTGGCGCTGGTCCGC 500
Oy 128 PheLeuPheTyIleAsnLeuTyGlySerIleLeuLeuLeuThrCysIleSerValHis 147
Db 501 TTCCTCTTCTATGCAACCTGCACGGCAGCATCTCTCTCCCTGACCTGCTCATCAGCTTCCAG 560
Oy 148 GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyArgThrArg----- 164
Db 561 CGCTACCTGGGCACTGTGCCACCGCTGGCC-----CCCTGGCACAAACGTGGGGGCGCT 614
Oy 165 -ArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValLeuGlnLeuLeuPr 184
Db 615 CGGGCTCGCTGGTGTGTGTGTAGCGGTGTGGTGGCGCTGACAAACCACTGCTGCC 674
Oy 184 oThrLeuAlaPheSerHisThrAspTyIleAsnGlyGlnMetIleTrpTyAspMetTh 204
Db 675 CACAGCCATCTTGGTGGCCACAGGATCCAGGATCAACCGCACTGCTGCTATGACCTCA 734
Oy 204 rSerGlnGluAsnPheAspArgLeuPheAlaTyGlyIleValLeuThrLeuSerGlyph 224
Db 735 CGCGCTCGCTG-GGACCCCACTATATGCCCTATGGCATGGCTCTCTCATCTGTCATCGCTT 793
Oy 224 eLeuSerLeu 227
Db 794 CTGGTGCCTT 803

RESULT 10
CNS05325/c
LOCUS      1101 bp      DNA      linear      GSS 26-JUL-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
            020W21 of library A from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION  AL318925
VERSION     AL318925.1 GI:9551809
KEYWORDS   GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
ORGANISM    Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
            1 (bases 1 to 1101)
REFERENCE   Roest Cröllius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissensbach,J.
            Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
            Nat. Genet. 25 (2), 235-238 (2000)
            20296633

TITLE
JOURNAL
MEDLINE

```

PUBMED 10835645  
REFERENCE 2 (bases 1 to 1101)  
AUTHORS Croillius,H.R., Jaillon,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.  
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis  
JOURNAL Genome Res. 10 (7), 939-949 (2000)  
MEDLINE 20359837  
PUBMED 10899143  
REFERENCE 3 (bases 1 to 1101)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (12-APR-2000)  
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.

## FEATURES

Source

1..1101  
/organism="Tetraodon nigroviridis"  
/db\_xref="taxon:99883"  
/clone="020M21"  
/clone\_lib="A"  
/note="Genoscope sequence ID : C0AA020AG11C1-end : T7"  
305 a 216 c 280 g 284 t 16 others

## BASE COUNT

## ORIGIN

Alignment Scores:  
Pred. No.: 1,73e-24 Length: 1101  
Score: 325.00 Matches: 67  
Percent Similarity: 62.50% Conservative: 38  
Best Local Similarity: 39.88% Mismatches: 61  
Query Match: 16.79% Indels: 3  
DB: 17 Gaps: 2

US-10-023-586B-2 (1-360) x CNS0532S (1-1101)

QY 43 CysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePhe 62  
Db 526 TGTCTGGCTTTCATCTCAACGCCAGCCCTGTGG---CTCTCTCGAGATGCGGCC 410  
QY 63 IleLeuGlyLeuProLeuAsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArg 82  
Db 466 GTCTTGGCTTTCATCTCAACGCCAGCCCTGTGG---CTCTCTCGAGATGCGGCC 410  
QY 83 TrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuTyrValLeu 102  
Db 409 TGAACCCCAACACCCCTCTCTTGTTCACCTCCAGCTTCTCTACATCTG 350  
QY 103 ---LeuProPheLeuIleIleThrTyrSerLeuAspArgTrpProPheGlyGluLeu 121  
Db 349 TCCTGCCACCCCTCATCTACTATTACGCCACCCAGCCGCTTC-GGCGTGGCA 291  
122 LeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 141  
Db 290 GCCTGCAAAATGTGGCTTCTCTTCTAGCTCAACCTCTACTGCAGCATCTTCTC 231  
QY 142 ThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyr 161  
Db 230 ACCTGCATCAGCGTGTACCGTTACGTGGGCATCTGCCACCCGATTAAGCGCTGA 171  
QY 162 ArgThrArgHisAlaTrpLeuGlyThrSerThrThrAlaLeuValValLeuGln 181  
Db 170 GTGAANCCCGANATTCCTACCTGGTGGCGCTCGGTGGGCTGTGTATAGTCTGT 111  
QY 182 LeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrpTyr 201  
Db 110 CTGGTGCCTTAACCTCATCTTGTCCACCATCTCCAGCGGGGACAAATATCNCCTGTCC 51  
QY 202 AspMetThrSerGlnGluAsnPhe 209  
Db 50 GATACCACNNCCCGGAGTATT 27

## RESULT 11

BI768397

## LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI768397 934 bp mRNA linear EST 25-SEP-2001  
603053844F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5203258 5',  
mRNA sequence.  
BI768397 GI:15759975  
EST.  
human.

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 934)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LAM11509 row: h column: 11  
High quality sequence stop: 765.

## FEATURES

source

Location/Qualifiers

1..934

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:5203258"

/clone\_lib="NIH\_MGC\_122"

/lab\_host="DH10B"

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;  
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source  
anonymous pool of 24 week female lung, 16 week female  
spleen, and 20-22 week male spleens. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.4 kb, insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 026. Note:  
this is a NIH\_MGC Library."

BASE COUNT 160 a 315 c 258 g 201 t

## ORIGIN

## Alignment Scores:

Pred. No.: 7,78e-24 Length: 934

Score: 318.00 Matches: 83

Percent Similarity: 52.13% Conservative: 27

Best Local Similarity: 39.34% Mismatches: 90

Query Match: 16.43% Indels: 13

DB: 13 Gaps: 4

US-10-023-586B-2 (1-360) x BI768397 (1-934)

QY 33 AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluLys 48  
Db 271 GACAAATGGCACAGCGCGCTGTGGCTTGGCCACCCACCCACCTGTCTACCGCGAGAA 330  
QY 49 TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu 68  
Db 331 TTCAGCAACTGCTGTGTATTCGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 390  
QY 69 AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr 88  
Db 391 AACATCTGTGTATTCACCATCTCCAGCTCCCGCGCGCGCTGTGTGTGTGTGTGTGT 450  
QY 89 TyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeu 107  
Db 451 TACACCCCTAAACCTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 510

QY 108 IleThrTyrSerLeuAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHis 127  
 Db 511 TACAACATGCCCAAGGTGATCACTGGCCCTTTGGCGACTTCGCTGCCGCCCTGGTCCGC 570  
 QY 128 PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis 147  
 Db 571 TTCTCTTCTATGCAACCTGCGCGGAGCATCTCTTCTCCTACCTGCATCAGCTTCACG 630  
 QY 148 GluPheLeuGlyValCysHisProLeuCysSer-LeuProTyrArgThrArgHisAl 167  
 Db 631 CGTACTGGGCACTCTCCACCCGCTGGCCCTGGCGCAACAGTGGGGCGCGGGCTGC 690  
 QY 167 atpLeuGlyThrSerThrThrThrAlaLeuValValLeuGlnLeuLeuProThrLeuAl 187  
 Db 691 CTGGTAGTGTGTAGCCGCTGGTGGCCCGTGACACACAGTGGCTGCCACAGCCAT 750  
 QY 187 aPheSerHisThrAspTyrIleAsn----GlyGlnMetIleTrpTyrAspMethrSerGl 206  
 Db 751 CTTGGCTGACACAGG-CATTCAAGCGTAAACCGAGCTGGTGTATGACCTCAGCCGCTC 809  
 QY 206 nGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeu-----ThrLeuSe 222  
 Db 810 TGG-CCTGGGCAACCACTATGCTGCTATGCAATGGCGCTCCACTGGTCACTGGTCTGGC 868  
 QY 222 rGlyPheLeuSerLeuLeuGlyHisPheGly 232  
 Db 869 TGCCTTTGCTGCCCTGTGGCTGCTACTGTC 899

RESULT 12  
 LOCUS BQ038875 623 bp mRNA linear EST 01-MAY-2002  
 DEFINITION pgnlc.pk010.113 normalized chicken lymphoid cDNA library Gallus  
 gallus cDNA clone pgnlc.pk010.113 5' similar to  
 SP|P34996|P2YR.CHICK P2Y PURINOCEPTOR 1 (ATP RECEPTOR) (P2Y1)  
 (PURINERGIC RECEPTOR) pir|S33733 G protein-coupled receptor  
 chicken emb|CA51176.1| (X73268) Atp receptor P2Y1 [Gallus gallus],  
 mRNA sequence.  
 ACCESSION BQ038875  
 VERSION BQ038875.2 GI:20383637  
 KEYWORDS EST.  
 SOURCE chicken.  
 ORGANISM Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 Phasianinae; Gallus.  
 1 (bases 1 to 623)  
 Morgan, R.W. and Burnside, J.  
 Chicken lymphoid ESTs  
 Unpublished (2001)  
 On Mar 27, 2002 this sequence version replaced gi:19772415.  
 Contact: Joan Burnside  
 Molecular Endocrinology  
 University of Delaware  
 40 Townsend Hall, Newark, DE 19717, USA  
 Tel.: 302 831-1345  
 Fax: 302-831-3411  
 Email: joan@udel.edu, www.chickest.udel.edu.

FEATURES  
 source  
 1..623  
 /organism="Gallus gallus"  
 /db\_xref="taxon:9031"  
 /clone="pgnlc.pk010.113"  
 /clone.lib="normalized chicken lymphoid cDNA library"  
 /sex="Male and Female"  
 /tissue\_type="thymus, bursa, spleen, PBL, bone marrow"  
 /lab\_host="E.coli EMDH10B"  
 /note="Vector: pCMVSPORT 6"  
 BASE COUNT 117 a 199 c 172 g 128 t  
 ORIGIN

Alignment Scores: 5.41e-24 Length: 623  
 Pred. No.:

Score: 317.00 Matches: 69  
 Percent Similarity: 58.60% Conservative: 23  
 Best Local Similarity: 43.95% Mismatches: 63  
 Query Match: 16.37% Indels: 2  
 Db: 14 Gaps: 2

US-10-023-586b-2 (1-360) x BQ038875 (1-623)

QY 53 TyrLeuSerLeuAlaTyrSerIlePheIleLeuGlyLeuProLeuAsnGlyThrVal 72  
 Db 135 TACTGCCCCACCTGCTACATCTCTGCTTCTATCACCAGGCTTCTGGGCACACAGCGTGGCC 194  
 QY 73 LeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThrTyrLeuValAsn 92  
 Db 195 ATCTGGATGTCGTCTTCCACATCGCGCGTGGAGCGGCATCTCGGTGTACATGTTTCAAC 254  
 QY 93 LeuMetValAlaAspLeuLeuTyrValLeu---LeuProPheLeuGlyIleLeuThrTyrSer 111  
 Db 255 CTGGCTCTGGCGCACTTCTGTATGTCTCAGCTGCGCGCCCTCATCTTCTACTACTTC 314  
 QY 112 LeuAspAspArgTrpProPheGlyGluLeuLeuCysLysLeuValHisPheLeuPheTyr 131  
 Db 315 AACAAACCCACTGGATCTTCGGGGAGCTCATGTGCAAGCTGCAGAGGTTCAATTTTCCAC 374  
 QY 132 IleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHisGlnPheLeuGly 151  
 Db 375 GTGAACCTCTACGCGCAGCATCTTCTCAGCTGCATAAGCGTGCACAGGTACACGGGC 434  
 QY 152 ValCysHisProLeuCysSerLeuProTyrArgThrArgArgHisAlaTrpLeuGlyThr 171  
 Db 435 GTCTGTCACCGCTGAAGTCTGCTGGGAGGCTGAAGAAGAAGACCGCTGACGTGACG 494  
 QY 172 SerThrThrAlaLeuValValLeuGlnLeuProThrLeuAlaPheSerHisThr 191  
 Db 495 TCCTGTGTGGCGCCCTGGTGGCGCTCATCGCGCCCATCTCTTCTACTCAGGACG 554  
 QY 192 AspTyrIleAsnGlyGlnMetIle---TrpTyrAspMethrSerGlnGlu 207  
 Db 555 NNNGTAGGAGGAACAAACCATCACGTGCTACGACACACCGCGNGACGACG 605

RESULT 13  
 LOCUS AL520218 936 bp mRNA linear EST 13-FEB-2001  
 DEFINITION AL520218 LTI\_NFL004\_NBC2 Homo sapiens cDNA clone CS0DB006VE03 5  
 prime, mRNA sequence.  
 ACCESSION AL520218  
 VERSION AL520218.1 GI:12783711  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 936)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES  
 source  
 1..936  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DB006VE03"  
 /clone.lib="LTI\_NFL004\_NBC2"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-stranded cDNA was digested with Not I and  
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com>

161 a 314 c 258 q 196 t 7 others

BASE COUNT

BASE COUNT	161 a	314 c	258 g	196 t	7 others
ORIGIN					

Alignment Scores:	
Pred. No.:	2.13e-23
Score:	314.00
Percent Similarity:	55.06%
Best Local Similarity:	44.94%
Query Match:	16.22%
DB:	9
Length:	936
Matches:	71
Conservative:	16
Mismatches:	65
Indels:	7
Gaps:	3

US-10-023-586B-2 (1-360) x AL520218 (1-936)

QY	33	AspMetAsnThrSerGlnGluGlnGlyLeu-----CysGlnPheSerGluIlys	48
Db	415	GACAATGGCACAGCCACGCCTTGGCGTTGCCACCACCAACCATCTGTCTATCCGCGAGAAC	474
QY	49	TyTLySglnValTyLeuSerLeuAlaTySerIleIlePheIleLeuGlyLeuProLeu	68
		:::     :::                 :::	
QY	475	TTCAACACTGCTGCTGCCACCTGTGTATTGGCGGTCTGGCGCTGGCTGGCTGGCGTG	534
QY	69	AsnGlyThrValLeuTrpHisSerTrpGlyGlnThrLysArgTrpSerCysAlaThrThr	88
		:::     :::     :::     :::     :::     :::     :::	
Db	535	AACATCTGTGCATTACCACGATCTGCACGTCCCGCGGGCCCTGACCCGACGCCCGG	594
QY	89	TyrLeuValAsnLeuMetValAlaaspLeuTyrValLeu---LeuProPheLeulle	107
		:::	
Db	595	TACACCTTAACACTTCTCTGGCTGACCTGCTATATGCTTCCTGCTGCCCTGGCTGCATC	654
QY	108	IleThr-Tyr-SerLeuAspArgTrpProPheGlyGlulLeuLeuCystylsLeuValHis	127
		:::                         :::     :::     :::     :::	
Db	655	TACAACTATGCCAACGGTGATCTACCTGGCCCTTTGGGCACTTGGCTGCGCTGGTCCGC	714
QY	128	PheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuThrCysIleSerValHis	147
Db	715	TTCTCTCTTATGCCAACCCTGCACGCACGATCTCTTCTCACCTGCATCAGCTTCAG	774
QY	148	GlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgArgHis	166
		::: :::	
Db	775	CGTACTTGGGATCTGCCACCCGCTGGCCCCCTGGCACAAACGTGGGGSGCCCGGGCT	834
QY	167	AlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeuGlnLeuLeuPro	184
Db	835	GCCTGCTAGTGTGTAGCCGTGGTGGCGCGGTGAM-AMCAGTAGCTCGCCA	887

RESULT 14  
AL549441  
AL549441 LTI\_NFL006\_P12 Homo sapiens cDNA clone CSODI044YA15 5  
prime, mRNA sequence.  
801 bp mRNA linear EST 16-FEB-2001  
INITIATION

ACCESSION	AL549441	
VERSION	AL549441.1	GI:12885426
KEYWORDS	EST.	
SOURCE	human	

ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 801)
AUTHORS	Li, W. B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope

FEATURES  
source 1. .801  
Location/Qualifiers  
Email: seq@genoscope.cns.fr, web : www.genoscope.cns.fr

RESULT 15

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="CS0DIO4YA15"  
/clone\_lib="LTI\_NFL006\_PL2"  
/tissue\_type="placenta"  
/note="Vector: pCWSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCWSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 616 8371 Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL : <http://fulllength.invitrogen.com>"

BASE COUNT	148 a	270 c	191 g	167 t	25 others
ORIGIN					

ORIGIN

Alignment Scores:		
Pred. No.:	4 58e-23	801
Score:	310.00	Matches: 76
Percent Similarity:	50.74%	Conservative: 27
Best Local Similarity:	37.44%	Mismatches: 90
Query Match:	16.01%	Indels: 10
DB:	9	Gaps: 3
US-10-023-586B-2 (1-360)	x	AL549441 (1-801)

US-10-023-586B-2 (1-360) x AL549441 (1-801)

Qy	33	AspMetAsnThrSerGlnGlnGlnLeu-----CysGlnPheSerGluLys	48
Db	200	GACATGGCAGCAGCCAGGCTGTGGCGTTGCCACCCACCACTGTCTTACATCCGCGAGAAC	259
Qy	49	TyrLysGlnValTyrLeuSerLeuAlaTyrSerIleIlePheIleLeuGlyLeuProLeu	68
Db	260	TTCCAGCACTGCTGCTGCMACCTGTATTCGGCGGTCTGGCGGCTGGCTGGCGCGTG	319
Qy	69	Asn-GlyThrValLeuTrpHisSerTrpGlyGlnThrLysAsrTgtrPser-CysAlaThrT	88
Db	320	AACATCTAGTGTCTATPACCCAGATCTGCAGCTCCCGCGCGGCCCTGACCCGACGCGCGCA	379
Qy	88	hrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyrValLeu---LeuProPhe-Leu	106
Db	380	TGTACACCTATATACCAAGAGTGATCACTGGCCCTTTGGCGACTTCGCCTCGCGMMTGTT	439
Qy	107	IleIleThrTyrSerLeuAsp-AspargtrpproPheGlyGluLeuLeuCysLysLeuVa	126
Db	440	ATCTACAACCTATGACCAAGAGTGATCACTGGCCCTTTGGCGACTTCGCCTCGCGMMTGTT	499
Qy	126	LHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeuLeuLeuThrCysIleSerVa	146
Db	500	MCGMTTCMTCTTCATATGCAAAACATGACGCGCAGCATCTMTCTWAMMCACMTGCATCAGTCT	559
Qy	146	LHisGlnPheLeuGlyValCysHisProLeuCysSerLeuProTyrArg---ThrArgAr	165
Db	560	CCAGCGCTACCTGGGGATCTGCCACCGCTGGCCCTGGCACAACACGTGGGGGCGCGCG	619
Qy	165	gHisAlaTrpLeuGlyThrSerThrTrpAlaLeuValValLeuGlnLeuLeuProThr	185
Db	620	GGCTGCTGGCTAGTGTGTGTASCCGTGGTGGCGGTGACAAACCCAGTKYCTGCCAC	679
Qy	185	rLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrrpTyrAspMetThrSe	205
Db	680	AGCCATCTTCGTGGTCCACAGGCATCCAGCGCTAACCGCACTGTCTGTGTAGCTTCAGCCCC	739
Qy	205	rGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuLeuLeuSerGlyPheLe	225
Db	740	GCCTGCCCTGGCCACCACTATATATGCGCTATGGCATGGCTCTCACTGTCTATCGGCTTCT	799
Qy	225	u 225	
Db	800	g 800	
RESULT 15			







GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:45:35 ; Search time 37 Seconds  
(without alignments)  
1296.493 Million cell updates/sec

Title: US-10-023-586b-2  
Perfect score: 1936  
Sequence: 1 MSLILLPSGRSGSRGAL.....QNKLGHPAGRKRCFGLNRS 360

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
A_Geneseq_101002:*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1798	92.9	333	22	AAE04125 Human G protein-co
2	1542	79.6	295	23	AAU11251 Human P2Y-like G p
3	1185	61.2	339	23	AAE18640 Human G-protein.co
4	921	47.6	170	22	AAE0935 Human nGPR12... Ho
5	421	21.7	374	22	AAE04390 Turkey P2Y nucleot
6	416	21.5	537	23	AAU74538 Human P2Y purinoc
7	395	20.4	328	17	AAE01224 Mouse pancreas G-p
8	394.5	20.4	377	22	AAE01144 Human purinergic r
9	392.5	20.3	328	17	AAE01225 Human placenta G-p
10	387.5	20.0	328	22	AAE04393 Human P2-purinergi

11	387.5	20.0	377	22	AAE04392 Human P2-purinergi
12	387.5	20.0	377	22	AAE01143 Human purinergic r
13	371	19.2	373	23	AAE20604 Mus musculus GPCR
14	370.5	19.1	375	16	AAE72457 Human P20 receptor
15	370	19.1	97	22	AAU25585 Human G Protein-Co
16	368.5	19.0	328	18	AAW09433 Human placenta pur
17	367	19.0	365	18	AAW23606 Human P2Y4 recepto
18	367	19.0	365	22	AAE04391 Human P2Y4 pyrimid
19	329.5	17.0	259	21	AAE45375 Gene 37 human secr
20	329	17.0	373	23	AAU10585 Purinergic recepto
21	328	16.9	373	22	AAE04389 Human P2-purinergi
22	328	16.9	373	23	AAU10583 Purinergic recepto
23	328	16.9	373	23	AAU10584 Human nGPR54 #2.
24	281.5	14.5	336	22	AAE0971 Human G-protein co
25	278.5	14.4	337	22	AAU04375 Human G-protein co
26	278.5	14.4	337	22	AAU04584 Human G-protein co
27	278.5	14.4	337	23	AAU15399 Human G-protein co
28	278.5	14.4	337	23	ABH81902 Human G-protein co
29	278.5	14.4	337	23	ABH83819 Human P2Y-like rec
30	278.5	14.4	337	23	ABE21803 Human AXOR89 (G-pr
31	278.5	14.4	337	23	ABH79438 Human P2Y1-li. Ho
32	278.5	14.4	337	23	AAU77600 Human P2Y1-like G
33	278.5	14.4	337	23	AAU14027 Human purinergic-r
34	278.5	14.4	337	23	AAE16171 Human G-protein co
35	278	14.4	230	22	AAW79297 Human protein SEQ
36	273.5	14.1	344	17	AAW04247 Human G-protein co
37	273.5	14.1	344	21	AAE15413 Human G-protein co
38	273.5	14.1	344	21	AAE67356 Human G-protein co
39	264.5	13.7	370	23	ABP61510 Human NF-kB activa
40	258.5	13.4	370	19	AAW62597 Human 7-transmembr
41	258.5	13.4	370	23	ABP61511 Human NF-kB activa
42	250.5	12.9	374	19	AAW51406 Human protease-act
43	249	12.9	339	15	AAE53752 Seven transmembran
44	249	12.9	339	18	AAE07617 Human G-protein th
45	249	12.9	339	19	AAW48733 Human R12 seven tr

ALIGNMENTS

RESULT 1  
AAG64125  
ID AAG64125 standard; Protein; 333 AA.  
XX  
AC AAG64125;  
XX  
DT 25-SEP-2001 (first entry)  
XX  
Human G protein-coupled receptor GPR71.  
XX  
Human: guanosine triphosphate binding protein-coupled receptor;  
KW G protein-coupled receptor; GPR8; GPR12; GPR16; GPR21; GPR40;  
KW GPR47; GPR51; GPR71; GPR72; cancer; liver cirrhosis;  
KW Alzheimer's disease; cytostatic; hepatotropic; nontropic;  
KW neuroprotective; gene therapy; peptide therapy.  
XX  
OS Homo sapiens.  
XX  
PN WO200148188-A1.  
XX  
PD 05-JUL-2001.  
XX  
PF 28-DEC-2000; 2000WO-JP09408.  
XX  
PR 28-DEC-1999; 99JP-0375152.  
XX  
PR 31-MAR-2000; 2000JP-0101339.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;  
PI Sugiyama T, Kishimoto T, Kanzaki K, Yasuda S, Inoue Y;  
XX  
XX WPI; 2001-425662/45.

DR N-PSDB; AAH73516.  
XX New DNA encoding guanosine triphosphate binding protein coupled  
PT receptors and their expression products for screening potential  
PT anticancer and neurotropic drugs and in diagnosis of these diseases  
XX  
PS Example 1; Page 132-135; 170pp; Japanese.  
XX  
XX The invention relates to nine human guanosine triphosphate binding  
CC protein (G protein)-coupled receptors designated GPRv8, GPRv12, GPRv16,  
CC GPRv21, GPRv40, GPRv47, GPRv51, GPRv71 and GPRv72, and to the  
CC genes encoding them. These genes and proteins and antibodies against  
CC the protein are useful in the treatment, prevention, diagnosis and  
CC investigation of diseases associated with G protein-coupled receptors,  
CC including cancer, cirrhosis of the liver and Alzheimer's disease.  
CC The present sequence is a G protein-coupled receptor of the invention.  
XX  
XX Sequence 333 AA;

Query Match 92.9%; Score 1798; DB 22; Length 333;  
Best Local Similarity 99.7%; Pred. No. 3.8e-191;  
Matches 331; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 29 MEKVDNMTSQEGLCQFSEKQVLSLAYSLIIFLGLPLNGTVLWHSWGOTKRWSCATT 88  
1 MEKVDNMTSQEGLCQFSEKQVLSLAYSLIIFLGLPLNGTVLWHSWGOTKRWSCATT 60  
Qy 89 YLVNLMVADLLVLLPFLIITYSLDDRPFGELCKLVHFLFYINLYGSILLTTCISVHQ 148  
61 YLVNLMVADLLVLLPFLIITYSLDDRPFGELCKLVHFLFYINLYGSILLTTCISVHQ 120  
Qy 149 FLGVCHPLCSPYRTRHAWLGSTTVALVQLLPTLAFSHTDYINGOMIWDMTSOEN 208  
121 FLGVCHPLCSPYRTRHAWLGSTTVALVQLLPTLAFSHTDYINGOMIWDMTSOEN 180  
Qy 209 FDLRFAYGIVLTLSGFLSLGHFGVLTGDPDQARCEPHEDRQHSQVHPDPTGVW 268  
181 FDLRFAYGIVLTLSGFLSLGHFGVLTGDPDQARCEPHEDRQHSQVHPDPTGVW 240  
Qy 269 PLHPLFCALPYHSLLLPHLLLSAFSGLPALDQSCGGLQDMASGECEQ 328  
241 PLHPLFCALPYHSLLLPHLLLSAFSGLPALDQSCGGLQDMASGECEQ 300  
Qy 329 GKNRVRLQLKRNKLGEPHAGRCPCGLNRS 360  
301 GKNRVRLQLKRNKLGEPHAGRCPCGLNRS 332

RESULT 2  
AAU11251  
ID AAU11251 standard; Protein; 295 AA.

XX AAU11251;

XX 26-FEB-2002 (first entry)

XX Human P2Y-like G protein-coupled receptor.

DE Human; P2Y-like G protein-coupled receptor; GPCR; COPD;  
XX chronic obstructive pulmonary disease; nervous system disease;  
KW Parkinson's disease; multiple sclerosis; dementia; stroke;  
KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;  
KW bacterial infection; fungal infection; protozoan infection;  
KW viral infection; pain; cancer; anorexia; bulimia; asthma;  
KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;  
KW angina pectoris; myocardial infarction; ulcer; inflammation; allergy;  
KW psychotic disorder; neurological disorder; anxiety; schizophrenia;  
KW manic depression; delirium; severe mental retardation; dyskinesia.

XX Homo sapiens.

OS W0200185764-A2.

XX

PD 15-NOV-2001.

XX 09-MAY-2001; 2001WO-EP05244.

XX 11-MAY-2000; 2000US-203582P.

PR 21-FEB-2001; 2001US-269857P.

XX (FARB ) BAYER AG.

XX Ramakrishnan S;

XX WPI; 2002-075242/10.

DR N-ESDB; AAS17746.

XX New polynucleotides for producing P2Y-like G protein-coupled receptors  
PT (GPCR) that are used for screening inhibitors or regulators of human  
PT P2Y-like GPCR, especially useful for treating pain, cancer or  
PT neurological disorders

XX Claim 25; Fig 2; 114pp; English.

XX The invention relates to an isolated polynucleotide encoding a P2Y-like  
CC G protein-coupled receptor (GPCR) polypeptide, its fragment,  
CC derivative or allele, a host cell containing an expression vector  
CC comprising the polynucleotide and screening for agents that regulate the  
CC GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR  
CC polypeptide, which may be employed for screening agents that inhibit or  
CC regulate human P2Y-like GPCR. The reagent or inhibitor of the human  
CC P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR  
CC disorders, particularly COPD (chronic obstructive pulmonary disease),  
CC peripheral or central nervous system disease (e.g. Parkinson's  
CC disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and  
CC many other diseases and disorders listed in the specification), benign  
CC prostatic hyperplasia or urinary incontinence. A pharmaceutical  
CC composition containing the modulators and/or regulators of P2Y-like  
CC GPCR is useful for modulating the activity of a P2Y-like GPCR.  
CC In particular, these are useful for treating, preventing or ameliorating  
CC infections (e.g. bacterial, fungal, protozoan or viral infections), pain,  
CC cancer, anorexia, bulimia, asthma, acute heart failure, hypotension,  
CC hypertension, osteoporosis, diabetes, angina pectoris, myocardial  
CC infarction, ulcers, inflammation, allergies, psychotic or neurological  
CC disorders (e.g. anxiety, schizophrenia, manic depression, delirium,  
CC severe mental retardation or dyskinesias). The present sequence is  
CC the P2Y-like GPCR of the invention.

XX Sequence 295 AA;

Query Match 79.6%; Score 1542; DB 23; Length 295;  
Best Local Similarity 98.6%; Pred. No. 9.6e-163;

Matches 284; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 29 MEKVDNMTSQEGLCQFSEKQVLSLAYSLIIFLGLPLNGTVLWHSWGOTKRWSCATT 88

Db 1 MEKVDNMTSQEGLCQFSEKQVLSLAYSLIIFLGLPLNGTVLWHSWGOTKRWSCATT 60

Qy 89 YLVNLMVADLLVLLPFLIITYSLDDRPFGELCKLVHFLFYINLYGSILLTTCISVHQ 148

Db 61 YLVNLMVADLLVLLPFLIITYSLDDRPFGELCKLVHFLFYINLYGSILLTTCISVHQ 120

Qy 149 FLGVCHPLCSPYRTRHAWLGSTTVALVQLLPTLAFSHTDYINGOMIWDMTSOEN 208

Db 121 FLGVCHPLCSPYRTRHAWLGSTTVALVQLLPTLAFSHTDYINGOMIWDMTSOEN 180

Qy 209 FDLRFAYGIVLTLSGFLSLGHFGVLTGDPDQARCEPHEDRQHSQVHPDPTGVW 268

Db 181 FDLRFAYGIVLTLSGFLSLGHFGVLTGDPDQARCEPHEDRQHSQVHPDPTGVW 240

Qy 269 PLHPLFCALPYHSLLLPHLLLSAFSGLPALDQSCGGLQDMASGECEQ 316

Db 241 PLHPLFCALPYHSLLLPHLLLSAFSGLPALDQSCGGLQDMASGECEQ 288

RESULT 3

AAE18640  
ID AAE18640 standard; Protein; 339 AA.  
XX  
AC AAE18640;  
XX  
DT 17-MAY-2002 (first entry)  
XX  
DE Human G-protein coupled receptor (GCREC-1).  
XX  
KW Human; G-protein coupled receptor; GCREC-1; cell proliferative disorder;  
KW neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;  
KW metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;  
KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;  
KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;  
KW osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;  
KW transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;  
KW neurotropic; neuroprotective; cardiant; immunosuppressive; anorectic;  
KW virucide; receptor.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 89..109  
FT /note="Transmembrane domain"  
FT Domain 130..149  
FT /note="Transmembrane domain"  
XX  
XX W020210387-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 25-JUL-2001; 2001WO-US23433.  
XX  
XX 27-JUL-2000; 2000US-221478P.  
XX 03-AUG-2000; 2000US-223268P.  
XX 21-AUG-2000; 2000US-227054P.  
XX 08-SEP-2000; 2000US-231121P.  
XX 13-SEP-2000; 2000US-232243P.  
XX 15-SEP-2000; 2000US-232691P.  
XX 22-SEP-2000; 2000US-235146P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;  
XX Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walia NK;  
XX Hafalia AJA, Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L;  
XX Graul RC, Warren BA, Lee EA, Ding L;  
XX  
XX WPI; 2002-188744/24.  
XX N-PSDB; AAD29667.  
XX  
XX New human G-protein coupled receptor polypeptide for diagnosis,  
XX prevention and treatment of cell proliferative, neurological,  
XX cardiovascular, gastrointestinal, autoimmune/inflammatory, and  
XX metabolic disorders -  
XX  
XX Claim 1; Page 114-115; 150pp; English.  
XX  
XX The invention relates to novel human G-protein coupled receptors  
XX (GCREC) and their encoding polynucleotides. GCREC is useful as an  
XX immunogen for preparing monoclonal and polyclonal antibodies. GCREC is  
XX useful for diagnosing, treating and preventing a cell proliferative  
XX disorder (e.g., hepatitis, psoriasis, cancer), a neurological disorder  
XX (e.g., epilepsy, Alzheimer's disease, Pick's disease, Huntington's  
XX disease, Parkinson's disease), a cardiovascular disorder (e.g.,  
XX atherosclerosis, hypertension, myocardial infarction), gastrointestinal  
XX disorder (e.g., gastritis, cirrhosis, Crohn's disease), an autoimmune/  
XX inflammatory disorder (e.g., acquired immunodeficiency syndrome (AIDS),  
XX allergy, anaemia, asthma, rheumatoid arthritis), a metabolic disorder  
XX (e.g., diabetes, obesity, osteoporosis), and viral infections. GCREC is  
XX useful in a number of drug screening techniques, and to analyse the  
XX proteome of a tissue or cell type. GCREC is useful for creating knockin

CC humanised animals or transgenic animals to model human diseases, in  
CC somatic or germline gene therapy, to generate a transcript image of a  
CC tissue or cell type, for detecting differences in the chromosomal  
CC location due to translocation, inversion, etc., among normal, carrier  
CC or affected individuals, and as hybridization probes for mapping  
CC naturally occurring genomic sequences. GCREC is useful in Southern or  
CC northern analysis, dot blot or other membrane-based technologies, in PCR  
CC technologies, in dipstick, pin, multiformat enzyme linked immunosorbant  
CC (ELISA)-like assays, and in microarrays utilising fluids or tissues from  
CC patients to detect altered GCREC expression. The present sequence is  
CC human GCREC-1.  
XX  
SQ Sequence 339 AA;  
XX  
XX Query Match 61.2%; Score 1185; DB 23; Length 339;  
XX Best Local Similarity 99.6%; Pred. No. 6.1e-123;  
XX Matches 225; Conservative 0; Mismatches 1; Indels: 0; Gaps 0;  
QY 1 MSLILPSRSGSRGSRGALLLEGASRDMKXDMNTSQEGLQCFSEKYKQVLSLAYSI 60  
DB 1 MSLILPSRSGSRGSRGALLLEGASRDMKXDMNTSQEGLQCFSEKYKQVLSLAYSI 60  
QY 61 IFILGLPLNGTVLWHSWGQTKRNSCATTYLVNLMVADLLYVLLPFLIITYSLDDRWPFGE 120  
DB 61 IFILGLPLNGTVLWHSWGQTKRNSCATTYLVNLMVADLLYVLLPFLIITYSLDDRWPFGE 120  
QY 121 LKCLVHFLFYINLYGSITLLTCTISVHOFGLGVCHPLCPSLPTRRHAWLGSTTWALVYL 180  
DB 121 LKCLVHFLFYINLYGSITLLTCTISVHOFGLGVCHPLCPSLPTRRHAWLGSTTWALVYL 180  
QY 181 QLLPTLAFSHPTDYINGQMIWYDMTSQENFDRLFAYGIVLTLSGFLS 226  
DB 181 QLLPTLAFSHPTDYINGQMIWYDMTSQENFDRLFAYGIVLTLSGFLS 226  
RESULT 4  
AAG80935  
ID AAG80935 standard; Protein; 170 AA.  
XX  
AC AAG80935;  
XX  
DT 28-AUG-2001 (first entry)  
XX  
DE Human ngPCR12.  
XX  
KW G protein-coupled receptor; ngPCR; seven transmembrane receptor;  
KW signal transduction; schizophrenia; thyroid disorder; renal failure;  
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;  
KW cardiovascular disease; proliferative disorder; hormonal disorder;  
KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;  
KW attention deficit-hyperactivity disorder/attention deficit disorder;  
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;  
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
KW neuroprotective.  
XX  
OS Homo sapiens.  
XX  
XX W0200136473-A2.  
XX  
XX 25-MAY-2001.  
XX  
XX 16-NOV-2000; 2000WO-US31581.  
XX  
XX 16-NOV-1999; 99US-0165838.  
XX 17-NOV-1999; 99US-0166071.  
XX 19-NOV-1999; 99US-0166678.  
XX 22-DEC-1999; 99US-0173396.  
XX 28-FEB-2000; 2000US-0184129.  
XX 28-FEB-2000; 2000US-0185421.  
XX 28-FEB-2000; 2000US-0185554.  
XX 02-MAR-2000; 2000US-0186530.  
XX 03-MAR-2000; 2000US-0186811.  
XX 09-MAR-2000; 2000US-0188114.





CC used to screen agonists and antagonists that modulate G-PCR  
XX activity, to raise antibodies and to develop assay systems.

SQ Sequence 328 AA;

Query Match 20.4%; Score 395; DB 17; Length 328;  
Best Local Similarity 31.9%; Pred. No. 4.4e-35;  
Matches 106; Conservative 55; Mismatches 137; Indels 34; Gaps 9;

QY 29 MEKVDNMTSQEGL-----COFSEKQKVYLSLAYSIIFILGLPLNGTVMHWSGOTKRW 84  
Db 1 MEQ-DNGTIQAPLPTTCVYREDKRLLLTPVYSVILVWGLPLNICVIAQICASRRLLT 59  
QY 85 CATTYLVNLMVADLYVL-LPFLIITYSLDDRWPFGEGLCKLVHFLFYINLYGSILLTLC 143  
Db 60 RSAVYTLNALADLMYACSLPLLIYNYARGDHPFDGLACRFVRFYANLHGSILFLTC 119  
QY 144 ISVHQFLGVCHPLCSLPR-TRRHAWLGTSTWALVVLQLLPTAFSTHTDINGQMIYD 202  
Db 120 ISFQYLGICHPLASWHRGGRRAAVVGVVWLVATAQAQLPTAVFAATGTQRNRTCYD 179  
QY 203 MTSQENFORLAYGIVLTLGSLF-----SLGHHFGVLETDGQEPDQARGEPHEDROHSPSQ 258  
Db 180 LSPPTLSRYLPYGMALIVIGFLPFIALLACYCRMARRLCRDGCPAGPVAGERRSKAAR 239  
259 VHPDHTGWPVLPFCALPYHSLLLPHLL-SAFSGLPALDGSQCGLODMEASGECEQL 317  
Db 240 M-----AVVAAVFAISLPFHITKTAVLAVRSTPGVSCPVLFTFAAAYKGT 287  
QY 318 PQPS-----PVLSF-----KGGKNNRVLLQKL 339  
Db 288 PFASVNSVLDPIFYFTOOKFRROPHDLLQRL 319

RESULT 8  
AAE01144  
ID AAE01144 standard; Protein; 377 AA.

XX AAE01144;

DT 17-JUL-2001 (first entry)

XX Human purinergic receptor P2Y2 protein.

XX Human; purinergic receptor; P2Y2; antidepressant; vulnary; hypotensive;  
KW anticonvulsant; antiinflammatory; therapy; nervous system disorder;  
KW epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;  
KW irritable bowel disorder; reproductive system disorder; hypertension;  
KW peripheral vascular disease; immune system disorder; chronic bronchitis;  
KW premature ejaculation; asthma; neuromuscular disease.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 258 /note= "Encoded by TFC"

FT Misc-difference 334 /note= "Wild type Arg substituted with Cys"

FT Misc-difference 350 /note= "Wild type Gly substituted with Glu"

FT Misc-difference 359 /note= "Wild type Phe substituted with Ser"

XX US6214581-B1.

XX 10-APR-2001.

XX 13-NOV-1998; 98US-0191136.

XX 16-JAN-1998; 98US-0071298.

XX 16-JAN-1998; 98US-0071669.

XX 16-JAN-1998; 98US-0008185.

XX 16-JAN-1998; 98US-0008526.

XX

PA (ABBO ) ABBOTT LAB.

PI Lynch KJ, Burgard EC, Van Biesen T;

XX WPI; 2001-315459/33.

DR N-PSDB; AAD04981.

XX Novel isolated polynucleotide encoding human purinergic P2X3 receptor  
polypeptide useful for identifying potentially therapeutic compounds  
that modulate or otherwise interact with P2X containing receptors -

PS Example 14; Fig 12B; 53pp; English.

XX The present sequence is human P2Y2 receptor protein. P2X  
receptors are ligand-gated ion channels while P2Y receptors operate  
generally through a G-protein coupled system. P2X purinoreceptor drugs  
are potential therapeutic agents in several disorders including central  
nervous system or peripheral nervous system conditions, e.g., epilepsy,  
pain, depression, neurodegenerative disorders, disorders of the skeletal  
muscle such as neuromuscular diseases, disorders of the reproductive  
system, asthma, peripheral vascular disease, hypertension, immune system  
disorders, irritable bowel disorder, premature ejaculation, cystic  
fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the activity  
of extracellular nucleotide triphosphates to regulate chloride secretion  
in human airway epithelia.

XX Sequence 377 AA;

Query Match 20.4%; Score 394.5; DB 22; Length 377;

Best Local Similarity 32.8%; Pred. No. 6.1e-35;

Matches 117; Conservative 40; Mismatches 121; Indels 79; Gaps 12;

QY 43 COFSEKQKVYLSLAYSIIFILGLPLNGTVMHWSGOTKRWSCATTYLVNLMVADLYVL 102  
Db 25 CRENEDFKYVLLPVSYGVVGVGLCLNAVALYIFLCRLKTNASTYWFHLAYSALYAA 84  
QY 103 -LPFLIITYSLDDRWPFGEGLCKLVHFLFYINLYGSILLTLCISVHQFLGVCHPLCSLPY 161  
Db 85 SLPLLVYVYARGDHPFSTVLCCKLVRFELTYNLYGSILFLTCISVHRCGLVRLRLRW 144  
QY 162 RTRRHAWLGTSTWALVVLQLLPTLAFSTHTDYNQMTWYDTSOENFDRLFAYGIVLTL 221  
Db 145 GRARTARVAGAVVVLVLAQAPVLYFVTTARGGRVCHTDSAPELFSRFVAYSSVM-- 202  
QY 222 SGFSLGLHFGVLF-----TDGQEPDQARGEPHEDROHSPSQVHP 261  
Db 203 ---LGLL--FAVPPAVILVCYVLMARRLLKPAYGTSG-----GLPRAKRKSVRT--- 246  
QY 262 DHPTGVWPLHPLF--CALPYH-----SLLPHILLSAFS-----GLPALDGSQCG 304  
Db 247 -----IAYVLAVFALCFPFHVTFTLYYSFRSLDLSCHTLNAINMAYKVTRPLASANC- 300  
QY 305 LQDMEASGECEQLPQSPVLSFKGKNNRVLLQKLQKNGKLGEPHAGKRCRP-GLNRS 360  
Db 301 -----LDPVLIFLAGQRLVRFARDAKP-PTGSPATPARCRLGLRRS 341

RESULT 9

AAE01225

ID AAE01225 standard; Protein; 328 AA.

XX AAE01225;

AC 26-AUG-1996 (first entry)

DT Human placenta G-protein coupled receptor protein.

XX G-protein coupled receptor protein; G-PCR; agonist; antagonist;

KW cystic fibrosis; incontinence; diabetes; diagnosis; therapy.

XX Homo sapiens.







DR WPI; 2001-315459/33.  
XX N-PSDB; AAD04980.  
XX Novel isolated polynucleotide encoding human purinergic P2X3 receptor  
PT polypeptide useful for identifying potentially therapeutic compounds  
PT that modulate or otherwise interact with P2X containing receptors  
XX Example 14; Fig 11B; 53pp; English.  
XX The present sequence is human wild type P2Y2 receptor protein, P2X  
CC receptors are ligand-gated ion channels while P2Y receptors operate  
CC generally through a G-protein coupled system. P2X purinoreceptor drugs  
CC are potential therapeutic agents in several disorders including central  
CC nervous system or peripheral nervous system conditions, e.g., epilepsy,  
CC pain, depression, neurodegenerative disorders, disorders of the skeletal  
CC muscle such as neuromuscular diseases, disorders of the reproductive  
CC system, asthma, peripheral vascular disease, hypertension, immune system  
CC disorders, irritable bowel disorder, premature ejaculation, cystic  
CC fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the activity  
CC of extracellular nucleotide triphosphates to regulate chloride secretion  
CC in human airway epithelia.  
XX Sequence 377 AA;  
SQ  
Query Match 20.0%; Score 387.5; DB 22; Length 377;  
Best Local Similarity 32.7%; Pred. No. 3.7e-34;  
Matches 117; Conservative 41; Mismatches 119; Indels 81; Gaps 13;  
QY 43 CQFSEKQVYLSLAYSIIFILGLPLNGTVLHWSGOTKRWSCATTYLVNLMVADLLVYL 102  
DB 25 CRFNEDEKYYLLPVSVGVVGLCLNVALYIFLCRLKTNWASTTMYHFLAVSDLYAA 84  
QY 103 -LPFLITYSLDDRWPFGEGLCKLVHFLYINLYGTSILLTCTISVHOFGLGVCHPLCSLPY 161  
DB 85 SLPLYYIYARGDHFNFSTVCLVKLRFELFTNLYCSILFCTISVHRCGLVLRPLSLRW 144  
QY 162 RTRHAWLGTSWTWALVQLLPTLAFSTHDYINGOMIWDYMTSQENFDRFLFAYGIVLTL 221  
DB 145 GRARYARRVAGAVWLVLACQAPVLVFTVTSARGRVTCCHDTSAPELFSRFVAYSSVM-- 202  
QY 222 SGFLSLILGHGVLF-----TDCQEPDQARGEPHEDRQHSPOVHP 261  
DB 203 ---LGLL--FAVFAVLVCYVLMARRLLKPAYTSG-----GLPRAKRKSVRT-- 246  
QY 262 DHTGTWVPLHPLF--CALPVH-----SLLPHHLLSAFS-----GLPALDGSQC 304  
DB 247 -----TAVVLAVFALCFELPHVTRTYLSPRSLLDLSCHTLNAINMAYKVTPLASANS- 300  
QY 305 LQDMEASGECEQLPQSPVLSFKGKNRVLQKLRQNKLGHPA--GRKRCFGLNRS 360  
DB 301 -----LDPVLYFLAGORLVRFARDKP-PTGPSPATPARRL-GLRRS 341  
RESULT 13  
AAE20604  
ID AAE20604 standard; Protein; 373 AA.  
XX AC AAE20604;  
XX 01-JUL-2002 (first entry)  
XX Mus musculus GPCR ATP-receptor P2U1 protein.  
DE Mouse; G-protein coupled receptor; GPCR; transgenic animal; receptor;  
KW genetic therapy; pharmacological; genetic disease; neuropsychological;  
KW neurological; psychotic illness; nephrotropic; gynaecological;  
KW psychostimulant; ATP-receptor P2U1.  
XX Mus musculus.  
XX WO200203793-A2.  
XX 17-JAN-2002.

XX 10-JUL-2001; 2001WO-US21923.  
XX 10-JUL-2000; 2000US-217058P.  
PR 10-JUL-2000; 2000US-217179P.  
PR 10-JUL-2000; 2000US-217223P.  
PR 10-JUL-2000; 2000US-217253P.  
PR 10-JUL-2000; 2000US-217255P.  
PR 10-JUL-2000; 2000US-217256P.  
PR 10-JUL-2000; 2000US-217257P.  
PR 11-JUL-2000; 2000US-217347P.  
PR 11-JUL-2000; 2000US-217629P.  
PR 12-JUL-2000; 2000US-217537P.  
PR 12-JUL-2000; 2000US-218069P.  
PR 12-JUL-2000; 2000US-218074P.  
PR 12-JUL-2000; 2000US-218358P.  
PR 27-JUL-2000; 2000US-221483P.  
PR 07-AUG-2000; 2000US-223120P.  
PR 07-AUG-2000; 2000US-223122P.  
PR 26-OCT-2000; 2000US-243958P.  
PR 15-NOV-2000; 2000US-249408P.  
PR 20-NOV-2000; 2000US-252299P.  
PR 16-JAN-2001; 2001US-262113P.  
PR 16-JAN-2001; 2001US-262205P.  
XX (DELT-) DELTAGEN INC.  
XX Allen KD, Brennan TJ;  
XX WPI; 2002-164574/21.  
DR N-PSDB; AAD32937.  
XX Novel non-human transgenic animal, especially transgenic mice useful  
PT for identifying an agent that modulates expression or function of  
PT target gene, comprises disruptions in target G protein-coupled receptor  
PT gene  
XX Disclosure; Fig 19; 101pp; English.  
XX The invention relates to a non-human transgenic animal having targeted  
CC G-protein coupled receptor (GPCR) gene disruptions in melanocortin-3 gene  
CC 5-Hr-2B gene, chemokine receptor 9A gene, glucocorticoid-induced receptor  
CC gene, orphan GPR10 (UHR)-1 gene, orphan GPR14 gene, orphan GPR15 gene,  
CC beta chemokine receptor (E01) gene, endothelial differentiation GPCR3  
CC (EDG3) gene, Atp receptor P2U1 gene or adenosine 3 receptor gene. The  
CC transgenic animal is useful for identifying an agent that modulates the  
CC expression or function of the target gene, for identifying an agent that  
CC ameliorates a phenotype associated with a disruption in the target gene.  
CC A transgenic construct is useful for producing a transgenic animal,  
CC preferably a transgenic mouse. The transgenic animal is useful for  
CC testing the efficacy of proposed genetic and pharmacological therapies  
CC for human genetic diseases, such as neurological, neuropsychological or  
CC psychotic illnesses. The transgenic animal is also useful as models for  
CC diseases, disorders or conditions associated with phenotypes relating to  
CC a disruption in a target, and to identify pharmaceuticals, therapies,  
CC drugs and interventions which may be effective in treating a disease or  
CC other phenotypic characteristics of the animal. An agent which modulates  
CC the expression of the target gene is useful as a therapeutic for treating  
CC conditions associated with a disruption of the target gene. The present  
CC sequence is mouse GPCR ATP-receptor P2U1 protein.  
XX Sequence 373 AA;  
SQ  
Query Match 19.2%; Score 371; DB 23; Length 373;  
Best Local Similarity 31.6%; Pred. No. 2.5e-32;  
Matches 113; Conservative 42; Mismatches 115; Indels 88; Gaps 12;  
QY 43 CQFSEKQVYLSLAYSIIFILGLPLNGTVLHWSGOTKRWSCATTYLVNLMVADLLVYL 102  
DB 25 CRFNEDEKYYLLPVSVGVVGLCLNVALYIFLCRLKTNWASTTMYHFLAVSDLYAA 84  
QY 103 -LPFLITYSLDDRWPFGEGLCKLVHFLYINLYGTSILLTCTISVHOFGLGVCHPLCSLPY 161  
DB 85 SLPLYYIYARGDHFNFSTVCLVKLRFELFTNLYCSILFCTISVHRCGLVLRPLSLRW 144  
QY 162 RTRHAWLGTSWTWALVQLLPTLAFSTHDYINGOMIWDYMTSQENFDRFLFAYGIVLTL 221  
DB 145 GRARYARRVAGAVWLVLACQAPVLVFTVTSARGRVTCCHDTSAPELFSRFVAYSSVM-- 202  
QY 222 SGFLSLILGHGVLF-----TDCQEPDQARGEPHEDRQHSPOVHP 261  
DB 203 ---LGLL--FAVFAVLVCYVLMARRLLKPAYTSG-----GLPRAKRKSVRT-- 246  
QY 262 DHTGTWVPLHPLF--CALPVH-----SLLPHHLLSAFS-----GLPALDGSQC 304  
DB 247 -----TAVVLAVFALCFELPHVTRTYLSPRSLLDLSCHTLNAINMAYKVTPLASANS- 300  
QY 305 LQDMEASGECEQLPQSPVLSFKGKNRVLQKLRQNKLGHPA--GRKRCFGLNRS 360  
DB 301 -----LDPVLYFLAGORLVRFARDKP-PTGPSPATPARRL-GLRRS 341  
RESULT 13  
AAE20604  
ID AAE20604 standard; Protein; 373 AA.  
XX AC AAE20604;  
XX 01-JUL-2002 (first entry)  
XX Mus musculus GPCR ATP-receptor P2U1 protein.  
DE Mouse; G-protein coupled receptor; GPCR; transgenic animal; receptor;  
KW genetic therapy; pharmacological; genetic disease; neuropsychological;  
KW neurological; psychotic illness; nephrotropic; gynaecological;  
KW psychostimulant; ATP-receptor P2U1.  
XX Mus musculus.  
XX WO200203793-A2.  
XX 17-JAN-2002.

Db 85 SLPLLVYYARGDHPFSTVLCVKLVRFLFYTNLYCSILFLTCISVHRCLGVLRLPLHSLRW 144  
 Qy 162 RTRRHWLGSTTTWALVQLLPTLAFSHTDYINGOMIWMYDMTSOENFDRLFAYGIVLTL 221  
 Db 145 GRARVARRAAVWVVLVACQAPVLYFVTTSVRGFRITCHDTSARELSHFVAYSSVM-- 202  
 Qy 222 SGFLSLLGHFGVLF-----TDGQEPDQARGEHPEDRO--HSPSOV 259  
 Db 203 ---LGLL--FAVPFVILVCVLMARLLKPAYGTG-----GLPRAKRKSVRTIALV 250  
 Qy 260 HPDHTPTGVWPLHPLFCALPYH-----SLLPHHLLSAFS-----GLPALDGSOGC 304  
 Db 251 -----LAVFAL-----CFLPFHVTRTLTYSFRLDLSCHTLNAINMAYKITRPLASANS- 300  
 Qy 305 LODMEASGEQOLPOPSVLFKGGKNRVRLQKLR-----ONKLGHPAGR 351  
 Db 301 -----LDPVLYLAGORLVRFARDAKPTEPTPSQARRKGLGHRPNR 343

RESULT 14  
 AAR72457  
 ID AAR72457 standard; Protein; 375 AA.  
 XX  
 AC AAR72457;

29-NOV-1995 (first entry)

DE Human P20 receptor.  
 XX  
 KW Epithelial mucosa; mucus; cystic fibrosis; asthma;  
 KW chronic bronchitis.

OS Homo sapiens.

PN WO9510538-A.

PD 20-APR-1995.

XX 04-OCT-1994; 94WO-US11260.

PR 15-OCT-1993; 93US-0138137.

XX (UMOR ) UNIV MISSOURI.

PA (UNNC-) UNIV NORTH CAROLINA.

XX Boucher RC, Erb LJ, Harden TK, Lustig KD, Parr CE;

PI Sullivan DM, Turner JT, Weisman GA;

DR WPI; 1995-169967/22.

DR N-PSDB; AAQ88134.

XX DNA encoding human P20 receptor and null cells expressing the  
 XX receptors - for stimulating or inhibiting growth of cultures of  
 XX mammalian cells, and for treating diseases of airway epithelial  
 XX mucosa, e.g. asthma.

XX Disclosure; Page 34; 47pp; English.

XX The sequence is that of the human P20 receptor. The receptor may  
 CC be used to treat diseases of epithelial mucosal surfaces by  
 CC enhancing or inhibiting mucus prodn. It may be used to treat  
 CC cystic fibrosis, asthma and chronic bronchitis.

XX Sequence 375 AA;

Query Match 19.1%; Score 370.5; DB 16; Length 375;  
 Best Local Similarity 32.7%; Pred. No. 2.8e-32;  
 Matches 115; Conservative 42; Mismatches 124; Indels 71; Gaps 13;

Qy 43 CQFSEKQVLSLAYSIYIFILGLPLNGTVLWHSWGQTKRWSCATYLVNLMVADLLXVL 102  
 Db 25 CRFNEEDLAYVLLPVSYGVVVCVGLCLNAGVLYIFLCRLKTNWASTYMFHLAVSDALYAA 84

Qy 103 -LPFLIITYSLDDRPFGELCKLKHFLFYINLYGSILLTLCISVHQFLGVCHPLCLSPY 161  
 Db 85 SLPLLVYYARGDHPFSTVLCVKLVRFLFYTNLYCSILFLTCISVHRCLGVLRLPLHSLRW 144  
 Qy 162 RTRRHWLGSTTTWALVQLLPTLAFSHTDYINGOMIWMYDMTSOENFDRLFAYGIVLTL 221  
 Db 145 GRARVARRAAVWVVLVACQAPVLYFVTTSVRGFRITCHDTSARELSHFVAYSSVM-- 201  
 Qy 222 SGFLSLLGHFGVLF-----TDGQEPDQARGEHPEDROHSPSOVHP 261  
 Db 202 ---LGLL--FAVPFVILVCVLMARLLKPAYGTG-----GLPRAKRKSVRT--- 245  
 Qy 262 DHTPTGVWPLHPLFCALPYH-----SLLPHHLLSAFSGLPALDGSQGLQDME 309  
 Db 246 -----IAVVLAVFALCFLPFHVTRTLTYSFRLDLSCHTLNAIN-----MAYKVTRLA 293  
 Qy 310 ASGECEQLPOPSVLFKGGKNRVRLQKLRQKLGHEPA-GRKRCPCGLNRS 360  
 Db 294 SANS- ----LDPVLYLAGORLVRFARDAKP-PTGSPATPARRTLGLRRS 339

RESULT 15

AAU25585

ID AAU25585 standard; Protein; 97 AA.

XX AC AAU25585;

XX 18-DEC-2001 (first entry)

XX Human G Protein-Coupled Receptor (GPCR) polypeptide #32.

XX Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia;  
 KW attention deficit disorder; anxiety; depression; bipolar disorder;  
 KW neurological disorder; Huntington's disease; dementia; obesity; anorexia;  
 KW metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis;  
 KW type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;  
 KW cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV;  
 KW viral infection; immunostimulant; neuroleptic; nootropic; tranquiliser;  
 KW antidepressant; anorectic; gene therapy.

XX Homo sapiens.

XX WO200162797-A2.

XX 30-AUG-2001.

XX 23-FEB-2001; 2001WO-US05676.

XX 23-FEB-2000; 2000US-0184247.

XX 23-FEB-2000; 2000US-0184303.

XX 23-FEB-2000; 2000US-0184304.

XX 23-FEB-2000; 2000US-0184305.

XX 02-MAR-2000; 2000US-0184397.

XX 02-MAR-2000; 2000US-0186457.

XX 03-MAR-2000; 2000US-0186810.

XX 09-MAR-2000; 2000US-0188064.

XX 13-MAR-2000; 2000US-0188880.

XX 03-APR-2000; 2000US-0194344.

XX 23-JUN-2000; 2000US-0213861.

XX 11-JUL-2000; 2000US-0217369.

XX 11-JUL-2000; 2000US-0217370.

XX 14-JUL-2000; 2000US-0218337.

XX 20-JUL-2000; 2000US-0218492.

XX (PHAA ) PHARMACIA & UPJOHN CO.

XX Vogeli G, Wood LS, Parodi LA, Lind P;

XX WPI; 2001-570628/64.

XX N-PSDB; AAS42837.

PT New isolated nucleic acid encoding a new G-protein coupled receptor  
 PT polypeptide for detecting receptor modulators that can treat mental

PT disorders, such as schizophrenia, anxiety, depression, or obesity -  
XX  
PS Claim 35; Page 82; 279pp; English.  
XX  
CC Sequences AAU25554-AAU25516 represent human G-protein coupled receptor  
CC (GPCR) polypeptides of the invention. The proteins and their associated  
CC DNA sequences can be used to identify compounds which bind to GPCR  
CC polypeptides and in screening for compounds that modulate GPCR activity.  
CC By screening a human subject for the presence of mutations in GPCR DNA, a  
CC GPCR-related disorder or a genetic predisposition can be diagnosed. The  
CC sequences can also be used for treatment and prevention of mental  
CC disorders such as schizophrenia, attention deficit disorder, anxiety,  
CC depression, dementia and bipolar disorder, neurological disorders such as  
CC Huntington's disease, Parkinson's disease and Tourette's syndrome,  
CC metabolic disorders such as obesity, anorexia and type 2 diabetes,  
CC cardiovascular disorders such as thrombosis, myocardial infarction,  
CC cardiomyopathy and atherosclerosis, viral infections caused by HIV and  
CC cancers.

XX SQ Sequence 97 AA;

Query Match 19.1%; Score 370; DB 22; Length 97;  
Best Local Similarity 100.0%; Pred. No. 4.9e-33;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSLIPLPSRSGSRGRRGALLLEGASRDMEKVDNMTSQEGLCQFSEKVKQVYLSLAYS I 60  
Db 24 MSLIPLPSRSGSRGRRGALLLEGASRDMEKVDNMTSQEGLCQFSEKVKQVYLSLAYS I 83  
QY 61 IFILGLPLNGTVLW 74  
Db 84 IFILGLPLNGTVLW 97

Search completed: May 22, 2003, 12:49:44  
Job time : 39 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:49:05 ; Search time 16 Seconds  
(without alignments)  
662.016 Million cell updates/sec

Title: US-10-023-586b-2  
Perfect score: 1936  
Sequence: 1 MSLTLLPSRSGSRGAL.....QNKLGHPAGRKRCPLNRS 360

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	395	20.4	328	US-08-513-974B-39	Sequence 39, Appl
2	395	20.4	328	US-08-513-974B-371	Sequence 371, Appl
3	392.5	20.3	328	US-08-513-974B-56	Sequence 56, Appl
4	392.5	20.3	328	US-08-513-974B-380	Sequence 380, Appl
5	383.5	19.8	327	US-08-513-974B-372	Sequence 372, Appl
6	374.5	19.3	375	US-08-442-134A-2	Sequence 2, Appl
7	374.5	19.3	375	US-08-444-581B-2	Sequence 2, Appl
8	374.5	19.3	375	US-08-446-088A-2	Sequence 2, Appl
9	374.5	19.3	375	US-08-559-524A-3	Sequence 3, Appl
10	374.5	19.3	375	US-08-749-707-3	Sequence 3, Appl
11	369	19.1	373	US-08-513-974B-373	Sequence 373, Appl
12	368.5	19.0	328	US-08-459-046-2	Sequence 2, Appl
13	340	17.6	362	US-08-513-974B-374	Sequence 374, Appl
14	326.5	16.9	373	US-08-559-524A-4	Sequence 4, Appl
15	326.5	16.9	373	US-08-749-707-4	Sequence 4, Appl
16	274.5	14.2	302	US-08-467-948A-30	Sequence 30, Appl
17	274.5	14.1	344	US-08-467-947A-30	Sequence 30, Appl
18	273.5	14.1	344	US-08-467-948A-8	Sequence 8, Appl
19	273.5	14.1	344	US-08-467-947A-8	Sequence 8, Appl
20	258.5	13.4	370	US-08-781-250-2	Sequence 2, Appl
21	250.5	12.9	408	US-08-742-440A-6	Sequence 6, Appl
22	249	12.9	339	US-08-153-848-44	Sequence 44, Appl
23	249	12.9	339	US-08-812-871-3	Sequence 3, Appl
24	249	12.9	339	US-09-299-843A-44	Sequence 44, Appl
25	249	12.9	339	US-09-088-337B-44	Sequence 44, Appl
26	249	12.9	339	PCT-US93-11153-44	Sequence 44, Appl
27	249	12.9	339	PCT-US95-07180-2	Sequence 2, Appl

28	248.5	12.8	391	1	US-07-816-283-2	Sequence 2, Appl
29	248.5	12.8	391	1	US-08-417-103-2	Sequence 2, Appl
30	248.5	12.8	391	1	US-08-417-103-14	Sequence 14, Appl
31	246	12.7	347	4	US-08-405-271A-24	Sequence 24, Appl
32	246	12.7	395	1	US-08-097-938-2	Sequence 2, Appl
33	246	12.7	395	1	US-08-097-938-5	Sequence 5, Appl
34	246	12.7	395	1	US-08-476-000-2	Sequence 2, Appl
35	246	12.7	395	1	US-08-476-000-5	Sequence 5, Appl
36	246	12.7	395	1	US-08-472-840-2	Sequence 2, Appl
37	246	12.7	395	1	US-08-472-840-5	Sequence 5, Appl
38	246	12.7	395	2	US-08-476-976-2	Sequence 2, Appl
39	246	12.7	395	2	US-08-476-976-5	Sequence 5, Appl
40	246	12.7	395	3	US-08-474-410-2	Sequence 2, Appl
41	246	12.7	395	3	US-08-474-410-5	Sequence 5, Appl
42	246	12.7	395	4	US-08-486-673B-2	Sequence 2, Appl
43	246	12.7	395	4	US-08-486-673B-5	Sequence 5, Appl
44	246	12.7	399	1	US-08-476-000-61	Sequence 61, Appl
45	246	12.7	399	1	US-08-472-840-61	Sequence 61, Appl

ALIGNMENTS

RESULT 1  
US-08-513-974B-39  
; Sequence 39, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01599  
; FILING DATE: 10-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-093989  
; FILING DATE: 19-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-057186  
; FILING DATE: 16-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-007177  
; FILING DATE: 20-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-326611  
; FILING DATE: 28-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-270017  
; FILING DATE: 02-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236357

```
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236356
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189274
; FILING DATE: 11-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189273
; FILING DATE: 11-AUG-1945
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189272
; FILING DATE: 11-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 45753
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-513-974B-39

Query Match 20.4%; Score 395; DB 3; Length 328;
Best Local Similarity 31.9%; Pred. No. 2.3e-27;
Matches 106; Conservative 55; Mismatches 137; Indels 34; Gaps 9;

QY 29 MEKVDMTSQBOGL----COFSEKYKQVLSLAYSIIFILGLPLNGTVLWHSWGOTKRW 84
DB 1 MEQ-DNCTIQAPGLPPTTCYVREDFKRLITPVYVVLVGLPLNVCVIAQICASRRTLT 59
QY 85 CATTYLVNLMVADLLYVL-LPFLITYSLDDRWPFGLCKLVHFLVINYLSILLTC 143
DB 60 RSVYTLNLADLMYACSLPLLTYNARGDHPFGLACRFVFLFYANLHGSILF LTC 119
QY 144 ISVHOFVGVCHPLCSLYR-TRRWLGSTTWALVVLVLLPTLAFSTDYINGOMIWD 202
DB 120 ISFORYLGICHPLASWIKRGRRAAVVGCVLWLVAVTAQCLPTAVFAATGIQRNRTVCYD 179
QY 203 MTSQENEDRLFAYGIVLTLSGLF----SLGHGCVLFTDGOEPDQARGEPHEDRQHSQS 258
DB 180 LSPPISTRLPYGMALTVIGFLPFTALLACYCRMARRLCRODGPAGVPAQERRSKAR 239
QY 259 VHPDPTGVNPLHPLFCALPYHSLLLPHLL-SAFSGLPALDGSQCGLQDMESGEQOL 317
DB 240 M-----AVVVAVFAISFLPFHITKTAYLAVRSTPGVSCPVLETFAAAYKGR 287
QY 318 PQPS-----PVLSF----KGGKRNVRLLQKL 339
DB 288 PFASVNSVLPDILEYFTQQRFRPHQLLQRL 319

RESULT 2
US-08-513-974B-371
; Sequence 371, Application US/08513974B
; Patent No. 6114139
; GENERAL INFORMATION:
; APPLICANT: Hinuma, Shuji
; APPLICANT: Hosoya, Masaki
; APPLICANT: Fujii, Ryo
; APPLICANT: Ohtaki, Tetsuya
; APPLICANT: Fukusumi, Shoji
; APPLICANT: Ohgi, Kazuhito
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,
; NUMBER OF SEQUENCES: 380
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,974B
; FILING DATE: 14-SEP-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA: PCT/JP95/01599
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-093989
; FILING DATE: 19-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-057186
; FILING DATE: 16-MAR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-007177
; FILING DATE: 20-JAN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-326611
; FILING DATE: 28-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-270017
; FILING DATE: 02-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236357
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-236356
; FILING DATE: 30-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189274
; FILING DATE: 11-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189273
; FILING DATE: 11-AUG-1945
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-189272
; FILING DATE: 11-AUG-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S.
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 45753
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 371:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 328 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-513-974B-371

Query Match 20.4%; Score 395; DB 3; Length 328;
Best Local Similarity 31.9%; Pred. No. 2.3e-27;
Matches 106; Conservative 55; Mismatches 137; Indels 34; Gaps 9;

QY 29 MEKVDMTSQBOGL----COFSEKYKQVLSLAYSIIFILGLPLNGTVLWHSWGOTKRW 84
DB 1 MEQ-DNCTIQAPGLPPTTCYVREDFKRLITPVYVVLVGLPLNVCVIAQICASRRTLT 59
QY 85 CATTYLVNLMVADLLYVL-LPFLITYSLDDRWPFGLCKLVHFLVINYLSILLTC 143
DB 60 RSVYTLNLADLMYACSLPLLTYNARGDHPFGLACRFVFLFYANLHGSILF LTC 119
QY 144 ISVHOFVGVCHPLCSLYR-TRRWLGSTTWALVVLVLLPTLAFSTDYINGOMIWD 202
DB 120 ISFORYLGICHPLASWIKRGRRAAVVGCVLWLVAVTAQCLPTAVFAATGIQRNRTVCYD 179
QY 203 MTSQENEDRLFAYGIVLTLSGLF----SLGHGCVLFTDGOEPDQARGEPHEDRQHSQS 258
DB 180 LSPPISTRLPYGMALTVIGFLPFTALLACYCRMARRLCRODGPAGVPAQERRSKAR 239
QY 259 VHPDPTGVNPLHPLFCALPYHSLLLPHLL-SAFSGLPALDGSQCGLQDMESGEQOL 317
DB 240 M-----AVVVAVFAISFLPFHITKTAYLAVRSTPGVSCPVLETFAAAYKGR 287
QY 318 PQPS-----PVLSF----KGGKRNVRLLQKL 339
DB 288 PFASVNSVLPDILEYFTQQRFRPHQLLQRL 319
```



Db 60 RSAYVTLNLALADLACSLPLIYNYAGDHWPFGLACRFVRFYANLHGSILFLTC 119  
QY 144 ISVHQFLGVCHPLCSLPYR-TRRHAMLGSTTWALVVLQLLPTLAFSHTDYINGOMIWD 202  
Db 120 ISFORVLGICHPLASHKHGRGRRAAVVCVWVLATVACQLPTAVFAATGIQRNRVCVD 179  
QY 203 MTSQENFDRFAYGIVLTLSGFL-----SLIGHFGLVLTQDQEPDQARGEPHEDRQHSPPSQ 258  
Db 180 LSPILSTRVLPYGMALTVIGFLFELPFIALLACYCMARRLCRODGPAGVPAQERRSKAAR 239  
QY 259 VPHDHTGVWPHLPLFCALPYHSLLLPHLL-SAFSGLPALDGSQGLQDMEASGECEQL 317  
Db 240 M-----AVVVAFAISFLPHITKTAYLAVRSTPGVSCFVLETFFAAYKGR 287  
QY 318 PQPS-----PVLSE-----KGGKNRVRLLOKL 339  
288 PFASVNSVLDPLIFYFTQOKFRQPHDLQRL 319  
RESULT 3  
US-08-513-974B-56  
; Sequence 56, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01599  
; FILING DATE: 10-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-093989  
; FILING DATE: 19-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-057186  
; FILING DATE: 16-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-007177  
; FILING DATE: 20-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-326611  
; FILING DATE: 28-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-270017  
; FILING DATE: 02-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236357  
; FILING DATE: 30-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236356  
; FILING DATE: 30-SEP-1994

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189274  
; FILING DATE: 11-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189273  
; FILING DATE: 11-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189272  
; FILING DATE: 11-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Resnick, David S.  
; REGISTRATION NUMBER: 34,235  
; REFERENCE/DOCKET NUMBER: 45753  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 328 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-513-974B-56  
Query Match 20.3%; Score 392.5; DB 3; Length 328;  
Best Local Similarity 33.0%; Pred. No. 3.9e-27;  
Matches 109; Conservative 42; Mismatches 142; Indels 37; Gaps 9;  
QY 33 DMVTSQEGGL-----COFSEYKQVYLSLAYSIIIFGLPLNGTVLHWSQGTNRWSCATT 88  
Db 4 DNGTQGAALGPLPTTCVYRENFKOLLPPVYSAVLAAGLPNLICVITQICTSRRLATTAV 63  
QY 89 YLVNLMVADLLYVL-LPFLIITYSLDDRPFGELLCCKLVHFLFYINLYGSLTLLTCTSVH 147  
Db 64 YTLNLALADLLYACSLPLIYNYAAGDHWPFGLACRFVRFYANLHGSILFLTCISFQ 123  
QY 148 QFLGVCHPLCSLPYR-TRRHAMLGSTTWALVVLQLLPTLAFSHTDYINGOMIWDMTSQ 206  
Db 124 RYLGIHPLAPWHKGRGRRAAVVCVWVLATVACQLPTAVFAATGIQRNRVCYDLSP 183  
QY 207 ENFDRFAYGIVLTLSGFL-----SLIGHFGLVLTQDQEPDQARGEPHEDRQHSPPSQVHPD 262  
Db 184 ALATHYMPYGMALTVIGFLFELPFIALLACYCGLAC-----RLCRQDQGAEPVAVQ 231  
QY 263 HPTGVNPLHPLFCALPYHSLLLPHLL-SAFSGLPALDGSQGLQDMEASGECEQLPQP- 320  
Db 232 ERGKAARMVAVVAAAFASFLPHITKTAYLAVRSTPGVSCFVLETFFAAYKGRPF 289  
QY 321 -----SPVLSF-----KGGKNRVRLLOKL 339  
Db 290 ASANSVLDPLIFYFTQOKFRRRPHELLQKL 319  
RESULT 4  
US-08-513-974B-380  
; Sequence 380, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA

	Query Match	20.3%;	Score 392.5;	DB 3;	Length 328;
	Best Local Similarity	33.0%;	Pred. No. 3.9e-27;		
	Matches 109;	Conservative 42;	Mismatches 142;	Indels 37;	Gaps 9;
QY	33	DMNTSQEQL----	QFSEKYYKOVYLSLAYSIIFI	GLPLNGVNLVHSMGQTRKWS	SCATT 88
Db	4	DNGTGALGLPPTTCVY	RENFKQLLPPVYSAVLAAGL	PLNICVITQICTSRRAL	RTAV 63
QY	89	YLVNLMVADLLVYL-	LPFLIITYSLDDRPFGELCKL	VHFLFYINLYSGISILLT	CISVH 147
Db	64	YTLNLALADLLYACSL	PLLIINYAOGDHPFGDFACRL	VRFLEYANLHGSILFLT	CISFQ 123
QY	148	QFIGVCHPICSLPYR-	TRRHAWLGTTTVALVVLQ	LLPTLAFSHTDYINGQMT	YDWTQS 206

APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1945  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 372:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 327 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-513-974B-372

Query Match 19.8%; Score 383.5; DB 3; Length 327;  
Best Local Similarity 40.9%; Pred. No. 2.4e-26;  
Matches 83; Conservative 32; Mismatches 81; Indels 7; Gaps 4;  
QY 29 MEKVDNMTSQEQL-----CQFSEKYKQVYLSLAYSIIFILGLPLNGTIVLWHSWGQTKRWS 84  
DB 1 MEQ-DNGTQAPGLPTTCVYREDEKRLLLTPVSVVVGVLNLCVIAQICASRRTLT 59  
QY 85 CATTYLVNLMVADLLYL-LPFLIITYSLDDRPFGELCKLVHFLFYINLYGSIILLTLC 143  
DB 60 RSAYVTLNLAALADYACSLPLLIYNYARCDHWPFGLACRFYRFLFYANLHGSILFTC 119  
QY 144 ISVHQFLGVCHPLCSLPYR-TRRHAWLGTSTTVALVQLLPTLAFSHTDYINGQMIWYD 202  
DB 120 ISFQVLGICHPLASHKHKGRRAANVCGVWVLATVACLRATAVTAATGIQRNRYCYD 179  
QY 203 MTSQENFDRLFYGVITLSGFL 225  
DB 180 LSPPIILSTRYPYGMALTIVIGFL 202

RESULT 6  
US-08-442-134A-2  
Sequence 2, Application US/08442134A  
Patent No. 5596088  
GENERAL INFORMATION:  
APPLICANT: Boucher, Richard C.  
APPLICANT: Weisman, Gary A.  
APPLICANT: Turner, John T.  
APPLICANT: Harden, Thomas K.  
APPLICANT: Parr, Claude E.  
APPLICANT: Sullivan, Daniel M.  
APPLICANT: Erb, Laura  
APPLICANT: Lustig, Kevin D.  
TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and  
TITLE OF INVENTION: Null Cells Expressing P2U Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bell, Seltzer, Park & Gibson  
STREET: Post Office Drawer 34009  
CITY: Charlotte  
STATE: No. 5596088th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/442,134A  
FILING DATE: 16-MAY-1995

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sibley, Kenneth D.  
REGISTRATION NUMBER: 31,665  
REFERENCE/DOCKET NUMBER: 5470-71A  
TELEPHONE: 919-420-2200  
TELEFAX: 919-881-3175  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-442-134A-2  
Query Match 19.3%; Score 374.5; DB 1; Length 375;  
Best Local Similarity 32.7%; Pred. No. 1.8e-25;  
Matches 115; Conservative 43; Mismatches 123; Indels 71; Gaps 13;  
QY 43 CQFSEKYKQVYLSLAYSIIFILGLPLNGTIVLWHSWGQTKRWSGATTYLVNLMVADLLYL 102  
DB 25 CRNEDEKYVLLPVSVGVVCGVLCGLNAVGLYIFLCRLKTNASTTYMFLHAYSDALYAA 84  
QY 103 -LPFLIITYSLDDRPFGELCKLVHFLFYINLYGSIILLTLCISVHOFGLGVCHPLCSLPY 161  
DB 85 SLPLLVIYVYARGDHPFSTVCLCKLVRFLEYTNLYCSILFTLCISVHRCGLGVLRPLSLRW 144  
QY 162 RTRRHAWLGTSTTVALVQLLPTLAFSHTDYINGQMIWYDNTSQENFDRLFYGVITLT 221  
DB 145 GRARYARRVAGAVVWLVLACQAPVLYEVVTS-ARGPLTCHDTSAPELFSRFVAYSSYM-- 201  
QY 222 SGFLSLILGHFGVLF-----TDQEPDQARGEPHEDRQHSPOVHP 261  
DB 202 ---LGLL--FAVFAVILVYCVLMARLLKPAVGTSG-----GLPRAKKSVRT---- 245  
QY 262 DHTGTWVPLHPLF-CALPYH-----SLLPHHLLSAFSGLPALDGSQCGLODME 309  
DB 246 ----IAVLAIVFALCFPLPHVTITLYYSPRSLDLSCHTLNAIN-----MAYKVTRLA 293  
QY 310 ASGECEQLPQSPVLSFKGKNRVLQKLRQNKLGHPA-GRKRCPLNRS 360  
DB 294 SANSC-----LDPVLYLAGQLRVRFARDAKP-PTGPSPATPARRTGLRRS 339  
RESULT 7  
US-08-444-581B-2  
Sequence 2, Application US/08444581B  
Patent No. 5607836  
GENERAL INFORMATION:  
APPLICANT: Boucher, Richard C.  
APPLICANT: Weisman, Gary A.  
APPLICANT: Turner, John T.  
APPLICANT: Harden, Thomas K.  
APPLICANT: Parr, Claude E.  
APPLICANT: Sullivan, Daniel M.  
APPLICANT: Erb, Laura  
APPLICANT: Lustig, Kevin D.  
TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and  
TITLE OF INVENTION: Null Cells Expressing P2U Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bell, Seltzer, Park & Gibson  
STREET: Post Office Drawer 34009  
CITY: Charlotte  
STATE: No. 5607836th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444.581B
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/442,134
; FILING DATE: 16-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-71A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-444-581B-2

Query Match          19.3%; Score 374.5; DB 1; Length 375;
Best Local Similarity 32.7%; Pred. No. 1.8e-25;
Matches 115; Conservative 43; Mismatches 123; Indels 71; Gaps 13;

; 43 CFSEKQVYLSLAYSIIIFILGLPLNGTVLWHSGOTKRWSCATTYLVNLMVADLLYVL 102
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 25 CRNEDEFYVLLPSYGVVGVLCGLNAGLVIFLCRLKTNASTYMFHLAVSDALYAA 84
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 103 -LPFLITYSYDDRWPFGECLLKVHFLFYINLYGSIILLTICISVHQFLGVCHPLCSLPY 161
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 85 SLPLLVYIYARGDHWPFSTVLCCKLVRFYTNLYCSILFLTCISVHRCGLVLRPLRSRW 144
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 162 RTRRHAWLGTSTWALVQLLPTLAFSHTDYINGOMIYDWTQENFDRLFAFGIVLTL 221
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 145 GRARYARRVAGAVVVLVACQAPLVFYVTTTS-ARGPLTCHDTSAPELFSRFVAYSSVM-- 201
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 222 SGFLSLHGHFGVLF-----TDGQEPDQARGEHPEDHROHSPSOVHP 261
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 202 ---LGLL--FAVPPFAVILVCVLMARLLKPAYTSG-----GLPRAKRKSVT----- 245
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 262 DHPTGVWPLHPLF--CALPYH-----SLLPHLLHSAFSGLPALDGSQCGLQDME 309
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 246 -----IAVVLAVFALCFLPFHVTRTYLVSFRSLDLSCHTLNAIN-----MAYKVTRLA 293
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 310 ASGECEQLPQSPVLSFGKGNRVRLQLKRLQNKLGHPA-GRKRCPLNRS 360
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 294 SANSC-----LDPVLYFLAGQLRVRFARDAKP-PTGPPSPATPARTLGLRRS 339
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
US-08-446-088A-2
Sequence 2, Application US/08446088A
Patent No. 5691156
GENERAL INFORMATION:
APPLICANT: Boucher, Richard C.
APPLICANT: Weisman, Gary A.
APPLICANT: Turner, John T.
APPLICANT: Harden, Thomas K.
APPLICANT: Parr, Claude E.
APPLICANT: Sullivan, Daniel M.
APPLICANT: Erb, Laura
APPLICANT: Lustig, Kevin D.
TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and
TITLE OF INVENTION: Null Cells Expressing P2U Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bell, Seltzer, Park & Gibson
STREET: Post Office Drawer 34009
CITY: Charlotte
STATE: No. 5691156th Carolina
COUNTRY: USA

```

```

; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.088A
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenneth D. Sibley
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-71C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 375 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-446-088A-2

Query Match          19.3%; Score 374.5; DB 1; Length 375;
Best Local Similarity 32.7%; Pred. No. 1.8e-25;
Matches 115; Conservative 43; Mismatches 123; Indels 71; Gaps 13;

; 43 CFSEKQVYLSLAYSIIIFILGLPLNGTVLWHSGOTKRWSCATTYLVNLMVADLLYVL 102
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 25 CRNEDEFYVLLPSYGVVGVLCGLNAGLVIFLCRLKTNASTYMFHLAVSDALYAA 84
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 103 -LPFLITYSYDDRWPFGECLLKVHFLFYINLYGSIILLTICISVHQFLGVCHPLCSLPY 161
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 85 SLPLLVYIYARGDHWPFSTVLCCKLVRFYTNLYCSILFLTCISVHRCGLVLRPLRSRW 144
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 162 RTRRHAWLGTSTWALVQLLPTLAFSHTDYINGOMIYDWTQENFDRLFAFGIVLTL 221
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 145 GRARYARRVAGAVVVLVACQAPLVFYVTTTS-ARGPLTCHDTSAPELFSRFVAYSSVM-- 201
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 222 SGFLSLHGHFGVLF-----TDGQEPDQARGEHPEDHROHSPSOVHP 261
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 202 ---LGLL--FAVPPFAVILVCVLMARLLKPAYTSG-----GLPRAKRKSVT----- 245
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 262 DHPTGVWPLHPLF--CALPYH-----SLLPHLLHSAFSGLPALDGSQCGLQDME 309
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 246 -----IAVVLAVFALCFLPFHVTRTYLVSFRSLDLSCHTLNAIN-----MAYKVTRLA 293
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 310 ASGECEQLPQSPVLSFGKGNRVRLQLKRLQNKLGHPA-GRKRCPLNRS 360
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
; 294 SANSC-----LDPVLYFLAGQLRVRFARDAKP-PTGPPSPATPARTLGLRRS 339
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 9
US-08-559-524A-3
Sequence 3, Application US/08559524A
Patent No. 5871963
GENERAL INFORMATION:
APPLICANT: Conley, Pamela B.
APPLICANT: Jantzen, Hans-Michael
TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/559,524A  
FILING DATE: 15-NOV-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Reid G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 044481-5010-00-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-467-7000  
TELEFAX: 202-467-7176  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-559-524A-3

Query Match 19.3%; Score 374.5; DB 2; Length 375;  
Best Local Similarity 32.7%; Pred. No. 1.8e-25;  
Matches 115; Conservative 43; Mismatches 123; Indels 71; Gaps 13;  
QY 43 CFSEKIKOVYLSLAYSIIFILGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLYYL 102  
Db 25 CRNEDEKVLVPSYGVVCLNGLNAVGLYIFLCRLKTNASTYMFHLAVSDALYAA 84  
QY 103 -LPFLIITYSLDRWPFGEELCKLVHFLYINLYSGILLTICISVHQFGLVCHPLCSLPY 161  
Db 85 SLPLVYYVARGDHPFSTVLCVRLVFLYTNLYCSILFLTICISVHRCGLVRLRLSLRW 144  
QY 162 RTRRHAWLGTSTWALVQLLPTLAFSTHDYINGOMIWDTSQENFDRFLFAYGIVLTL 221  
Db 145 GRARVARRVAGVWLVLAQAPVLYFVTTS-ARGPLTCHDTSAPELFSRFVAYSSVM-- 201  
QY 222 SGFLSLILGHFGVLF-----TDGQEPDQARGPHEPDROHSPSOVHP 261  
Db 202 ---LGLL--FAVPFVILCYVLMARRLLKPAYGTSG-----GLPRAKRSVRT---- 245  
QY 262 DHTGTWVPLHPLF--CALPYH-----SLLPLHLLSAFSGPLALDGSQGLQDME 309  
Db 246 ----IAVVLAVFALCFPLFHVTRTYLSFRSLDLSCHTLNAIN-----NAYKVTRLA 293  
QY 310 ASGECEQLPQSPVLSFKGKNRVRLLQKLRONKLGHEPA-GRKRCPCGLNRS 360  
Db 294 SANSC-----LDPVLYFLAGQLRVRFARDAKP-PTGSPSPATPARRTLGLRRS 339

RESULT 10  
US-08-749-707-3  
Sequence 3, Application US/08749707  
Patent No. 6063582  
GENERAL INFORMATION:  
APPLICANT: Conley, Pamela B.  
APPLICANT: Jantzen, Hans-Michael  
TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
STREET: 1800 M Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036-5869  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/749,707  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Reid G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 044481-5010-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-467-7000  
TELEFAX: 202-467-7176  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 375 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-749-707-3  
Query Match 19.3%; Score 374.5; DB 3; Length 375;  
Best Local Similarity 32.7%; Pred. No. 1.8e-25;  
Matches 115; Conservative 43; Mismatches 123; Indels 71; Gaps 13;  
QY 43 CFSEKIKOVYLSLAYSIIFILGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLYYL 102  
Db 25 CRNEDEKVLVPSYGVVCLNGLNAVGLYIFLCRLKTNASTYMFHLAVSDALYAA 84  
QY 103 -LPFLIITYSLDRWPFGEELCKLVHFLYINLYSGILLTICISVHQFGLVCHPLCSLPY 161  
Db 85 SLPLVYYVARGDHPFSTVLCVRLVFLYTNLYCSILFLTICISVHRCGLVRLRLSLRW 144  
QY 162 RTRRHAWLGTSTWALVQLLPTLAFSTHDYINGOMIWDTSQENFDRFLFAYGIVLTL 221  
Db 145 GRARVARRVAGVWLVLAQAPVLYFVTTS-ARGPLTCHDTSAPELFSRFVAYSSVM-- 201  
QY 222 SGFLSLILGHFGVLF-----TDGQEPDQARGPHEPDROHSPSOVHP 261  
Db 202 ---LGLL--FAVPFVILCYVLMARRLLKPAYGTSG-----GLPRAKRSVRT---- 245  
QY 262 DHTGTWVPLHPLF--CALPYH-----SLLPLHLLSAFSGPLALDGSQGLQDME 309  
Db 246 ----IAVVLAVFALCFPLFHVTRTYLSFRSLDLSCHTLNAIN-----MAYKVTRLA 293  
QY 310 ASGECEQLPQSPVLSFKGKNRVRLLQKLRONKLGHEPA-GRKRCPCGLNRS 360  
Db 294 SANSC-----LDPVLYFLAGQLRVRFARDAKP-PTGSPSPATPARRTLGLRRS 339

RESULT 11  
US-08-513-974B-373  
Sequence 373, Application US/08513974B  
Patent No. 6114139  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
APPLICANT: Hosoya, Masaki  
APPLICANT: Fujii, Ryo  
APPLICANT: Ohtaki, Tetsuya  
APPLICANT: Fukusumi, Shoji  
APPLICANT: Ohgi, Kazuhiro  
TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
NUMBER OF SEQUENCES: 380  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA: US/08/513.974B

FILING DATE: 14-SEP-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA: PCT/JP95/01599

APPLICATION NUMBER: 10-AUG-1995

FILING DATE: 10-AUG-1995

PRIOR APPLICATION DATA: JP 7-093989

APPLICATION NUMBER: 19-AUG-1995

FILING DATE: 19-AUG-1995

PRIOR APPLICATION DATA: JP 7-057186

APPLICATION NUMBER: 16-MAR-1995

FILING DATE: 16-MAR-1995

PRIOR APPLICATION DATA: JP 7-007177

APPLICATION NUMBER: 20-JAN-1995

FILING DATE: 20-JAN-1995

PRIOR APPLICATION DATA: JP 6-326611

APPLICATION NUMBER: 28-DEC-1994

FILING DATE: 28-DEC-1994

PRIOR APPLICATION DATA: JP 6-270017

APPLICATION NUMBER: 02-NOV-1994

FILING DATE: 02-NOV-1994

PRIOR APPLICATION DATA: JP 6-236357

APPLICATION NUMBER: 30-SEP-1994

FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA: JP 6-236356

APPLICATION NUMBER: 30-SEP-1994

FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA: JP 6-189274

APPLICATION NUMBER: 11-AUG-1994

FILING DATE: 11-AUG-1994

PRIOR APPLICATION DATA: JP 6-189273

APPLICATION NUMBER: 11-AUG-1945

FILING DATE: 11-AUG-1945

PRIOR APPLICATION DATA: JP 6-189272

APPLICATION NUMBER: 11-AUG-1994

FILING DATE: 11-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Resnick, David S.

REGISTRATION NUMBER: 34,235

REFERENCE/DOCKET NUMBER: 45753

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 373:

SEQUENCE CHARACTERISTICS:

LENGTH: 373 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-513-974B-373

Query Match 19.1%; Score 369; DB 3; Length 373;

Best Local Similarity 31.6%; Pred. No. 5.4e-25;

Matches 113; Conservative 42; Mismatches 115; Indels 88; Gaps 12;

QY 43 CQSEKYKQVLSLAYSIITFLGLPLNGTVMHWSGOTKRWSCATTYLVNLMVADLLYL 102

Db 25 CRNEDEKRYLLPVSYGVCVGLCLNVALYIFLCRLKTNASTTWTFHLVNSDSLYAA 84

QY 103 -LPFLIITYSLDDRWPEGLCKLVHFLFYINELYSILLTCTISVHFLGCHPLCSLPY 161

Db 85 SLPLLVYVYARGDHWPESTVLCKLVRELFVTLNLYCSILFLTCTISVHRCGLVRLPLSLRW 144

QY 162 RTRHAWLGTSTTVALVQLLPLAFSHDYINGQMIWDMTSDQENFDRFLFAYGIVLTL 221

Db 145 GRARYARRAAVVMVLVLAQAPVLYFTTSVTRTCHDTSARELFSHFVAYSSVM-- 202

QY 222 SGFLSLGLHFGVLF-----TDQEQPDQARPEHEDRQ--HSPSQV 259

Db 203 --LGLL--FAVPFSVILVCVLMARLLKPAYCTTG-----GLPRKRKSVRTIALV 250

QY 260 HPDHPTGVWPLHPLFCALPYH-----SLLPLPHLLSAFS-----GLPALDGSQCQ 304

Db 251 -----LAVFAL-----CFLPHVHTRTLYYSFRSLDLSCHTLNAINMAYKITRPLASANS- 300

QY 305 LODMEASGECEQLPQSPVLSFKGKKNRVRLLQKLR-----QNKLGEPHAGR 351

Db 301 -----LDPVLYFLAGQRLVRFARDAKPPTEPTPSQARRKLGRLHPNR 343

RESULT 12

US-08-459-046-2

: Sequence 2, Application US/08459046

: Patent No. 6008039

: GENERAL INFORMATION:

: APPLICANT: Coleman, Roger

: APPLICANT: Au-Young, Janice

: APPLICANT: Stuart, Susan G.

: TITLE OF INVENTION: A No. 6008039el Human Purinergic P2U Receptor

: NUMBER OF SEQUENCES: 2

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

: STREET: 3330 Hillview Avenue

: CITY: Palo Alto

: STATE: CA

: COUNTRY: USA

: ZIP: 94304

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: PatentIn Release #1.0, Version #1.30

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/459,046

: FILING DATE:

: CLASSIFICATION: 435

: ATTORNEY/AGENT INFORMATION:

: NAME: Luther, Barbara J.

: REGISTRATION NUMBER: 33954

: REFERENCE/DOCKET NUMBER: PF-0038 US

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 415-855-0555

: TELEFAX: 415-852-0195

: INFORMATION FOR SEQ ID NO: 2:

: SEQUENCE CHARACTERISTICS:

: LENGTH: 328 amino acids

: TYPE: amino acid

: TOPOLOGY: linear

: MOLECULE TYPE: protein

US-08-459-046-2

Query Match 19.0%; Score 368.5; DB 3; Length 328;

Best Local Similarity 31.8%; Pred. No. 5.1e-25;

Matches 105; Conservative 42; Mismatches 146; Indels 37; Gaps 9;

QY 33 DMNFSQOGL-----CQSEKYKQVLSLAYSIITFLGLPLNGTVMHWSGOTKRWSCATT 88

Db 4 DNGTDQALGPLPTTCVYRENFQKLLPPVYSVAVLAPALNICVITQICTSRRLTRAV 63

QY 89 YLVNLMVADLLYL-LPFLIITYSLDDRWPEGLCKLVHFLFYINLYGSLLLTCTISVH 147

Db 64 YTNLALPDLUYACSLPLLYNYAQQDHWPEGFACRLVRFYANLHGRLLFTCTISFQ 123

QY 148 QFLGVCHPLCSLPYR-TRRHAWLGTSTTVALVQLLPLAFSHDYINGQMIWDMTSDQ 206

Db 124 RYLGLCHPLAPWKRGGRRAWLVCVAVLAVTTQCLPTAIFATGIQRNRTVCYDLSPP 183

QY 207 ENFDRFPAYGIVLTLSCFL-----SLIGHFGVLTDDGQEPQARGEHEDRQHSQVHPD 262

Db 184 ALATHYMPYGMALTIVIGFLPLFAALLACYCILLAC-----RLCRQDGPAPVPAQ 231

QY 263 HPTGVWPLHPLFCALPYHSLLLPHLL-SAFSGPLALDGSQCGLQDMEASGECEQLPQP- 320

Db 232 ERRGAARMVAAVAVGIGISFLPFHITKTAYLAVRSTPGVPCTV--LEAFAAAYKGRDPF 389

QY 321 -----SPVLSF-----KGGKNRVLLOKL 339

Db 290 ASANSVLDPILFYFTOKKFRRPHELLOKL 319

RESULT 13

US-08-513-974B-374

Sequence 374, Application US/08513974B

Patent No. 6114139

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji

APPLICANT: Hosoya, Masaki

APPLICANT: Fujii, Ryo

APPLICANT: Ohtaki, Tetsuya

APPLICANT: Fukusumi, Shoji

APPLICANT: Ongi, Kazuhiro

TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,

TITLE OF INVENTION: PRODUCTION, AND USE THEREOF

NUMBER OF SEQUENCES: 380

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/513,974B

FILING DATE: 14-SEP-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP95/01599

FILING DATE: 10-AUG-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-093989

FILING DATE: 19-AUG-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-057186

FILING DATE: 16-MAR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 7-007177

FILING DATE: 20-JAN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-326611

FILING DATE: 28-DEC-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-270017

FILING DATE: 02-NOV-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-236357

FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-236356

FILING DATE: 30-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-189274

FILING DATE: 11-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-189273

FILING DATE: 11-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-189272

FILING DATE: 11-AUG-1994

ATTORNEY/AGENT INFORMATION:

NAME: Resnick, David S.

REGISTRATION NUMBER: 34,235

REFERENCE/DOCKET NUMBER: 45753

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 374:

SEQUENCE CHARACTERISTICS:

LENGTH: 362 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-513-974B-374

Query Match 17.6%; Score 340; DB 3: Length 362;

Best Local Similarity 30.9%; Pred. No. 1.9e-22;

Matches 98; Conservative 49; Mismatches 128; Indels 42; Gaps 11;

QY 53 YLSAYSIIFILGLPLNGTVLHMSGQTKRWSGATYVNLVADLLYL-LPFLITYS 111

Db 42 YLPTVYILVFITGFLGNSVAIMWFVHMFPWSGISVYMFNLALADFLYVLTLPALIFYF 101

QY 112 LDDRWPFGELCKLVHFLVINYLSGSLILLTCTISVHQFLGVCHPCLSLPYTRRHAWLGT 171

Db 102 NKTDWIFGDVMCKLQRFIFHVNLYSLFLTCTISVHRYTGVVHPLKSLGRLKKNVYVS 161

QY 172 STTWALVQLPLTFLAFSHTDYINGQMI-WYDMTISOENFDRFAYGIVLTL----SGFLS 226

Db 162 SLVWALVAVIAPILFYSGIGVRRNKTITCYDTADEYLSRYFVYSNCTVFMFCIPFIV 221

QY 227 LLGHFGLFTDQEPDQARGEHPEDHROHSPSQVHPDHTGWPFLHPLFCALPYHSLLLPH 286

Db 222 ILGCYGLI-----VKALIYKDLNDSPLRRKSTVLYII--VLTVF-AVSY----LPF 265

QY 287 HLLSAFSGLPALD--GSGCGLD-----MEASGECEQLPQSPVLSFKGKN- 331

Db 266 HVMKTLNLRALDFQTPQMCFAFNKDYATYQVYTRGLASLNSC-----VDPILYFLAGDTF 320

QY 332 RVRLLQKLKONKLGHP 348

Db 321 RRLSRATRKSSRRSEP 337

RESULT 14

US-08-559-524A-4

Sequence 4, Application US/08559524A

Patent No. 5871963

GENERAL INFORMATION:

APPLICANT: Conley, Pamela B.

APPLICANT: Jantzen, Hans-Michael

TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP

STREET: 1800 M Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20036-5869

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/559,524A

FILING DATE: 15-NOV-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Adler, Reid G.

REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: 044481-5010-00-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-467-7000

TELEFAX: 202-467-7176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-559-524A-4

Query Match 16.9%; Score 326.5; DB 2; Length 373;  
Best Local Similarity 36.7%; Pred. No. 3.1e-21;  
Matches 77; Conservative 35; Mismatches 83; Indels 15; Gaps 4;  
QY 53 YLSLAYSIIFILGLPLNGTVLHWSGQTKRWSCATTVLVNLMVADLLYL-LPFLIITYS 111  
DB 53 YLPAYILVFIIGLGNVAIWVFMFKPWSGISVYMFNLALADFLYVLTLPALIFYF 112  
QY 112 LDDRWPFGELLCKLVHFLFYINLYGSILLTLCISVHQFLGVCHPLCSLPYTRRHAWLGT 171  
DB 113 NKTDWIFGDAMCKLQRFIFHVNLYGSILFLTLCISAHRYSGVWYPLKSLGRKKNAVYIS 172  
QY 172 STTWALVVLQLLPTLAESHDTYINGQMI-WYDMTSOENFDRLEAYGIVLTLSGF---LS 226  
DB 173 VLVWLIIVVVGISPLIFYSGTGIRKNKTITCYDTSDEYLSRYFIYSMCTTVAMFCVPLVL 232  
QY 227 LLGHFGVLTGQEPDQARGEPEHEDROHSP 256  
DB 233 ILGCYGLI-----VRALIYKDLNDSNP 253

RESULT 15  
US-08-749-707-4  
Sequence 4, Application US/08749707  
Patent No. 6063582  
GENERAL INFORMATION:  
APPLICANT: Conley, Pamela B.  
APPLICANT: Jantzen, Hans-Michael  
TITLE OF INVENTION: NOVEL PURINERGIC RECEPTOR  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
STREET: 1800 M Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036-5869  
COMPUTER READABLE FORM:  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/749,707  
FILING DATE: 15-NOV-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Adler, Reid G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 04481-5010-01-US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-467-7000  
TELEFAX: 202-467-7176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 373 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-749-707-4

Query Match 16.9%; Score 326.5; DB 3; Length 373;

Best Local Similarity 36.7%; Pred. No. 3.1e-21;  
Matches 77; Conservative 35; Mismatches 83; Indels 15; Gaps 4;  
QY 53 YLSLAYSIIFILGLPLNGTVLHWSGQTKRWSCATTVLVNLMVADLLYL-LPFLIITYS 111  
DB 53 YLPAYILVFIIGLGNVAIWVFMFKPWSGISVYMFNLALADFLYVLTLPALIFYF 112  
QY 112 LDDRWPFGELLCKLVHFLFYINLYGSILLTLCISVHQFLGVCHPLCSLPYTRRHAWLGT 171  
DB 113 NKTDWIFGDAMCKLQRFIFHVNLYGSILFLTLCISAHRYSGVWYPLKSLGRKKNAVYIS 172  
QY 172 STTWALVVLQLLPTLAESHDTYINGQMI-WYDMTSOENFDRLEAYGIVLTLSGF---LS 226  
DB 173 VLVWLIIVVVGISPLIFYSGTGIRKNKTITCYDTSDEYLSRYFIYSMCTTVAMFCVPLVL 232  
QY 227 LLGHFGVLTGQEPDQARGEPEHEDROHSP 256  
DB 233 ILGCYGLI-----VRALIYKDLNDSNP 253

Search completed: May 22, 2003, 12:51:36  
Job time : 17 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:49:50 ; Search time 22 Seconds  
(without alignments)  
1622.825 Million cell updates/sec

Title: US-10-023-586B-2  
Perfect score: 1936  
Sequence: 1 MLSTLLPSRSGSRGAL.....QNKLGHPAGKRCPCFLNRS 360

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2.6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2.6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2.6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2.6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2.6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2.6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2.6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2.6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2.6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2.6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2.6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2.6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2.6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2.6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1189	61.4	338	9	US-09-885-453-4
2	921	47.6	170	9	US-09-782-974C-14
3	371	19.2	341	9	US-10-270-587-3
4	370	19.1	97	9	US-09-791-932-92
5	367	19.0	365	9	US-09-077-173A-2
6	346	17.9	362	9	US-10-092-135-3
7	346	17.9	362	9	US-10-092-135-4
8	346	17.9	362	9	US-09-779-679-28
9	329.5	17.0	373	9	US-10-092-135-7
10	328	16.9	373	9	US-10-092-135-6
11	326.5	16.9	373	9	US-10-092-135-5
12	320	16.5	299	9	US-10-270-144-4
13	281.5	14.5	336	9	US-09-782-974C-86
14	278.5	14.4	337	9	US-10-023-775B-2
15	278.5	14.4	337	9	US-10-270-144-2
16	278.5	14.4	337	9	US-10-188-405-8
17	278.5	14.4	337	9	US-09-885-453-1
18	278.5	14.4	337	10	US-09-943-798-4
19	274.5	14.2	302	9	US-10-024-494-30

20	273.5	14.1	344	9	US-10-024-494-8	Sequence 8, Appli
21	250.5	12.9	367	9	US-09-828-478-6	Sequence 6, Appli
22	249	12.9	339	9	US-09-828-478-4	Sequence 4, Appli
23	249	12.9	339	10	US-09-848-889-12	Sequence 12, Appli
24	249	12.9	339	10	US-09-788-133-2	Sequence 2, Appli
25	248.5	12.8	391	9	US-09-990-940-20	Sequence 20, Appli
26	246	12.7	347	10	US-09-823-114-24	Sequence 24, Appli
27	242.5	12.5	317	9	US-09-891-138A-2	Sequence 2, Appli
28	242.5	12.5	359	9	US-10-094-417-10	Sequence 10, Appli
29	242.5	12.5	359	9	US-09-782-974C-76	Sequence 76, Appli
30	242.5	12.5	359	10	US-09-739-151-2	Sequence 2, Appli
31	242	12.5	359	9	US-09-867-915-3	Sequence 3, Appli
32	242	12.5	359	10	US-10-218-574-9	Sequence 9, Appli
33	242	12.5	359	10	US-09-867-569-9	Sequence 8, Appli
34	240.5	12.4	391	9	US-09-841-720-8	Sequence 1, Appli
35	239.5	12.4	359	9	US-10-190-469-1	Sequence 25, Appli
36	239.5	12.4	366	9	US-09-779-679-25	Sequence 1, Appli
37	236	12.2	795	10	US-09-254-783A-1	Sequence 1, Appli
38	236	12.2	795	12	US-10-152-058-1	Sequence 2, Appli
39	234.5	12.1	276	10	US-09-943-798-2	Sequence 2, Appli
40	234.5	12.1	362	9	US-09-850-948-2	Sequence 2, Appli
41	234	12.1	359	10	US-09-966-871-81	Sequence 81, Appli
42	234	12.1	359	12	US-10-039-645-81	Sequence 81, Appli
43	233	12.0	427	9	US-10-007-132-4	Sequence 4, Appli
44	233	12.0	427	9	US-09-771-287-4	Sequence 2, Appli
45	232.5	12.0	334	10	US-09-765-034-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-885-453-4  
; Sequence 4, Application US/09885453  
; Publication No. US20030088080A1  
; GENERAL INFORMATION:  
; APPLICANT: Communi, Didier  
; TITLE OF INVENTION: RECEPTOR GPCR10  
; FILE REFERENCE: 9409/2082  
; CURRENT APPLICATION NUMBER: US/09/885,453  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 09/885,453  
; PRIOR FILING DATE: 2001-06-21  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: Amino acid sequence  
; LOCATION: (1)..(338)  
; OTHER INFORMATION: GPCR10 amino acid sequence  
US-09-885-453-4

Query Match	61.4%	Score 1189;	DB 9;	Length 338;
Best Local Similarity	98.7%;	Pred. No. 2.4e-106;		
Matches	225;	Conservative	0;	Mismatches 3;
			Indels	0;
			Gaps	0;
QY	1	MLSTLLPSRSGSRGALLLEGASRDMEKVDNMTSQEQGLQFSEKVKQVYLSLAYS	60	
Db	1	MLSTLLPSRSGSRGALLLEGASRDMEKVDNMTSQEQGLQFSEKVKQVYLSLAYS	60	
QY	61	IFILGLPLNGTIVLHWSGQTKRWSCATTYLVNLMVADLLYVLLPFLITYSLDDRWPFGE	120	
Db	61	IFILGLPLNGTIVLHWSGQTKRWSCATTYLVNLMVADLLYVLLPFLITYSLDDRWPFGE	120	
QY	121	LLCKLVHFLFYINLYGSIILLITCISVHQFLGVCHPCLSLPYTRRHAWLGSTTVALVWL	180	
Db	121	LLCKLVHFLFYINLYGSIILLITCISVHQFLGVCHPCLSLPYTRRHAWLGSTTVALVWL	180	
QY	181	QLLPTLAFSHTDYINGOMIWDNMTSQENFDRLFAIGVILTLGFLSLL	228	
Db	181	QLLPTLAFSHTDYINGOMIWDNMTSQENFDRLFAIGVILTLGFLSLL	228	

Db 181 QLLPTLAFSHTDYINGOMIWDMTSQENFDRFLFAYGIVLTLSGFFPSL 228

RESULT 2

US-09-782-974C-14

; Sequence 14, Application US/09782974C

; Publication No. US20030082534A1

; GENERAL INFORMATION:

; APPLICANT: Vogeli, Gabriel

; APPLICANT: Lind, Peter

; APPLICANT: Wood, Linda S.

; APPLICANT: Parodi, Luis A.

; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor

; FILE REFERENCE: 411USPHRM311

; CURRENT APPLICATION NUMBER: US/09/782,974C

; CURRENT FILING DATE: 2002-09-04

; PRIOR APPLICATION NUMBER: 60/165,838

; PRIOR FILING DATE: 1999-11-16

; PRIOR APPLICATION NUMBER: 09/714,449

; PRIOR FILING DATE: 2000-11-16

; PRIOR APPLICATION NUMBER: 60/198,568

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: 60/166,071

; PRIOR FILING DATE: 1999-11-17

; PRIOR APPLICATION NUMBER: 60/166,678

; PRIOR FILING DATE: 1999-11-19

; PRIOR APPLICATION NUMBER: 60/173,396

; PRIOR FILING DATE: 1999-12-28

; PRIOR APPLICATION NUMBER: 60/184,129

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: 60/185,421

; PRIOR FILING DATE: 2000-02-28

; PRIOR APPLICATION NUMBER: 60/185,554

; PRIOR FILING DATE: 2000-02-28

; PRIOR APPLICATION NUMBER: 60/186,530

; PRIOR FILING DATE: 2000-03-02

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 192

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 14

; LENGTH: 170

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-782-974C-14

Query Match 47.6%; Score 921; DB 9; Length 170;

Best Local Similarity 100.0%; Pred. No. 6.2e-81;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 83 WSCATTYLVNLMVADLLVLLPFLIITYSLDDRWPFGEELCKLVHFLFYINLYGSILLLT 142

Db 1 WSCATTYLVNLMVADLLVLLPFLIITYSLDDRWPFGEELCKLVHFLFYINLYGSILLLT 60

143 CISVHQFLGVCHPLCSLYPRTTTHAWLGTSTWALVQLLPTLAFSHTDYINGOMIWD 202

Db 61 CISVHQFLGVCHPLCSLYPRTTTHAWLGTSTWALVQLLPTLAFSHTDYINGOMIWD 120

Qy 203 MTSQENFDRFLFAYGIVLTLSGFLSLGLGHFGVLTGQEPDQARGEPHEDR 252

Db 121 MTSQENFDRFLFAYGIVLTLSGFLSLGLGHFGVLTGQEPDQARGEPHEDR 170

RESULT 3

US-10-270-587-3

; Sequence 3, Application US/10270587

; Publication No. US20030054487A1

; GENERAL INFORMATION:

; APPLICANT: Li, Yi

; TITLE OF INVENTION: Human G-Protein Coupled Receptor

; FILE REFERENCE: PF217C2

; CURRENT APPLICATION NUMBER: US/10/270,587

; CURRENT FILING DATE: 2002-10-16

; PRIOR APPLICATION NUMBER: US 09/908,593

; PRIOR FILING DATE: 2001-07-20

; PRIOR APPLICATION NUMBER: US 08/781,456

; PRIOR FILING DATE: 1997-01-10

; PRIOR APPLICATION NUMBER: US 60/009,902

; PRIOR FILING DATE: 1996-01-11

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 3

; LENGTH: 341

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-270-587-3

Query Match

Best Local Similarity 31.6%; Pred. No. 1.5e-27;

Matches 113; Conservative 42; Mismatches 115; Indels 88; Gaps 12;

Qy 43 CQSEKYNQVYLSLAYSIIFILGPNLCTVLMHSGQTKRWCATTYLVNLMVADLLVYL 102

Db 22 CRNEDEKYLVPYSYGVCVGLCLNVVALYIFLCRLKTNASTYMFHLAVSDSLAA 31

Qy 103 -LPFLIITYSLDDRWPFGEELCKLVHFLFYINLYGSILLTLCISVHQFLGVCHPLCSLPY 161

Db 82 SLPLLVYYARGDHWPFSTVLCKLVRFYINLYCSILLFICISVHRCLGVLRPLHSLRW 141

Qy 162 RTRHAWLGTSTWALVQLLPTLAFSHTDYINGOMIWDYMTSQENFDRFLFAYGIVLT 221

Db 142 GRARYARRVAAMVVLVACQAPVLYFVTTSVRGTRITCHDTSARELFSHFVAYSSVM-- 199

Qy 222 SGFLSLGLGHFGVLP-----TDGEPDQARGEPHEDRQ--HSPQOV 259

Db 200 ---LGLL--FAVPFSEVLVCVLMARLLKPAYGTG-----GLPRAKRKSVRTIALV 247

Qy 260 HPDHPTGVMPLHPLFCALPYH-----SLILPHLLLSAFS-----GLPALDGSOCG 304

Db 248 -----LAVFAL---CFLPFHVTRTYYSFRSLDLSCHTLNAINMAYKITRPLASANS- 297

Qy 305 LODMEASGECEQLPQSPVLSFKGKKNRVLQLKLR-----ONKLGEPHAPGR 351

Db 298 -----LDPVLYELAGQLRVREARDAKPPTPTSPQARRKLGHLHRPNR 340

RESULT 4

US-09-791-932-92

; Sequence 92, Application US/09791932

; Publication No. US20030003451A1

; GENERAL INFORMATION:

; APPLICANT: Vogeli, Gabriel

; APPLICANT: Parodi, Luis A.

; APPLICANT: Hiebsch, Ronald R.

; APPLICANT: Lind, Peter

; APPLICANT: Kaytes, Paul S.

; APPLICANT: Ruff, Valerie

; APPLICANT: Huff, Rita M.

; APPLICANT: Wood, Linda S.

; TITLE OF INVENTION: No. US20030003451A1el G Protein-Coupled Receptors Cross-Ref

; FILE REFERENCE: 00325.US1

; CURRENT APPLICATION NUMBER: US/09/791,932

; CURRENT FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: 60/184,305

; PRIOR FILING DATE: 2000-02-23

; PRIOR APPLICATION NUMBER: 60/184,304

; PRIOR FILING DATE: 2000-02-23

; PRIOR APPLICATION NUMBER: 60/184,303

; PRIOR FILING DATE: 2000-02-23

; PRIOR APPLICATION NUMBER: 60/184,397

; PRIOR FILING DATE: 2000-02-23

; PRIOR APPLICATION NUMBER: 60/184,247

; PRIOR FILING DATE: 2000-02-23

; PRIOR APPLICATION NUMBER: 60/188,880

; PRIOR FILING DATE: 2000-03-13

; PRIOR APPLICATION NUMBER: 60/217,369

; PRIOR FILING DATE: 2000-07-11

PRIOR APPLICATION NUMBER: 60/217,370  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/218,492  
PRIOR FILING DATE: 2000-07-20  
PRIOR APPLICATION NUMBER: 60/186,810  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: 60/188,064  
PRIOR FILING DATE: 2000-03-09  
PRIOR APPLICATION NUMBER: 60/186,457  
PRIOR FILING DATE: 2000-03-02  
PRIOR APPLICATION NUMBER: 60/213,861  
PRIOR FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: 60/194,344  
PRIOR FILING DATE: 2000-04-03  
PRIOR APPLICATION NUMBER: 60/218,337  
PRIOR FILING DATE: 2000-07-14  
NUMBER OF SEQ ID NOS: 184  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 92  
LENGTH: 97  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-791-932-92

Query Match 19.1%; Score 370; DB 9; Length 97;  
Best Local Similarity 100.0%; Pred. No. 3.5e-28;  
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLIPLSRGSGSRGALLLEGASRDMKDYDMNTSQEGLCOFSEKYQVYLSAYSI 60  
DB 24 MSLIPLSRGSGSRGALLLEGASRDMKDYDMNTSQEGLCOFSEKYQVYLSAYSI 83  
QY 61 IFILGLPLNGTVLW 74  
DB 84 IFILGLPLNGTVLW 97

RESULT 5  
US-09-077-173A-2  
Sequence 2, Application US/09077173A  
Publication No. US20030082674A1  
GENERAL INFORMATION:  
APPLICANT: RECEPTOR AND NUCLEIC ACID MOLECULE ENCODING  
TITLE OF INVENTION: SAID RECEPTOR  
NUMBER OF SEQUENCES: 4  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0; Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/077,173A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/BE 96/00123  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-077-173A-2

Query Match 19.0%; Score 367; DB 9; Length 365;  
Best Local Similarity 41.1%; Pred. No. 3.9e-27;  
Matches 79; Conservative 30; Mismatches 75; Indels 8; Gaps 3;

QY 43 CFQSEKYQVYLSAYSIIFILGLPLNGTVLWHSWGOTKRWSCATTYLVNLMVADLLYL 102  
DB 27 CFWEDEPKFILLPVSAVVFVLGLGNAPTLWLFIFRLRPWDATATYMFHLAUSDLYVL 86.

QY 103 -LPFLIITYSLDDRWPFPGELLCKLVHFLFYINLYGSILLTCTISVHOFGLGVCHPLCLSPY 161  
DB 87 SLPTLIIVYAAAHNHWPFGTEICKFVRFELFYVNLVCSVLELTCISVHRYLGICHPFLRALRW 146  
QY 162 RTRHAWLGTSTTWALVVLQLLPLTAFSHTDYINGQMIWTDMTSQENFDRFLFAYGIVLTL 221  
DB 147 GRPLRAGLLCLAVNLVVGCLVPLNFVTTSNKGTTVLCHDTTRPEEFDHVHFES----- 201  
QY 222 SGFLSLILGHFGV 233  
DB 202 SAVMGLL--FGV 211

RESULT 6  
US-10-092-135-3  
Sequence 3, Application US/10092135  
Publication No. US20030054374A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR  
FILE REFERENCE: D0134.NP  
CURRENT APPLICATION NUMBER: US/10/092,135  
CURRENT FILING DATE: 2002-03-06  
PRIOR APPLICATION NUMBER: US 60/273,808  
PRIOR FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 60/278,983  
PRIOR FILING DATE: 2001-03-27  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 3  
LENGTH: 362  
TYPE: PRT  
ORGANISM: GALLUS GALLUS  
US-10-092-135-3

Query Match 17.9%; Score 346; DB 9; Length 362;  
Best Local Similarity 30.5%; Pred. No. 4.1e-25;  
Matches 97; Conservative 49; Mismatches 128; Indels 44; Gaps 10;

QY 53 YLSLAYSIIIFILGLPLNGTVLWHSWGOTKRWSCATTYLVNLMVADLLYL-LPFLIITYS 111  
DB 42 YLPTVYILVITGFLGNSVAIMFVFMHMPWSGISVYMFNLALADFLYLTLPALIIFYF 101  
QY 112 LDDRWPFGEELCKLVHFLFYINLYGSILLTCTISVHOFGLGVCHPLCLSPYRTRHAWLGT 171  
DB 102 NKTDWIFGDVWCKLQRFIFHVNLYGSILFVLCISVHRYTGVVHPLKSLGRLLKKNAYVS 161  
QY 172 STTWALVVLQLLPLTAFSHTDYINGQMI-WYDMTSQENFDRFLFAYGIVLTL-----SGFLS 226  
DB 162 SLVWALVAVIAPIILFYSGTGVRNKTITCYDTTADDEVLSRSFYVSMCTTVFMFCIPFIV 221  
QY 227 LLGHFGVLFITDGOEPDQARGEPHEDRHSQVHPDHTGVWPHLPFCALPYHSL-LLP 285  
DB 222 ILGCYGLI-----VKALIYKDLNLSPLRRK-----SYLVIIIVTVFAVSYLP 264  
QY 286 HLLLSAFSGPLALD---GSOCGLQD-----MEASGECEQLPQSPVLSFKGKN 331  
DB 265 FHVKNLTNLRARLDFQTPQCAFNDKVYATYQVTRGLASLNSC-----VDPILYFLAGDT 319  
QY 332 -RVRLQLKRLQNLGEHP 348  
DB 320 FRRRLSRATRKSSRRSEP 337

RESULT 7  
US-10-092-135-4  
Sequence 4, Application US/10092135  
Publication No. US20030054374A1  
GENERAL INFORMATION:  
APPLICANT: Bristol-Myers Squibb Company  
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR  
FILE REFERENCE: HGPRBMV27



Best Local Similarity 36.7%; Pred. No. 1.7e-23;  
Matches 77; Conservative 36; Mismatches 82; Indels 15; Gaps 4;  
QY 53 YLSLAYSIIIFILGLPLNGIVLHWSGQTKRWSCATTYLVNLMVADLLYVL-LPFLIITYS 111  
Db 53 YLPNAVILYFIIGFLGNSVAINMFVFMKPSGISVYMFNLALADFLYVLTLPALIFYF 112  
QY 112 LDDRWPFGELLCKLVHFLFYINLYGSILLTICISVHQFLGVCHPLGSLPYRTRHAWLGT 171  
Db 113 NKTDWIFGDVMCKLQRFIFHVNLVGSILFLTCISAHRYSGVVPYPLKSLGRKKNAIYVS 172  
QY 172 STTWALVVLQLLPTLAFSHTDY-INQMTIYDWTSONFDRFLPAYGIVLTLSGF-----LS 226  
Db 173 VLWLVVVAISPIIFYSGTGRKKNKTVCTYDSTDEYLRSYFIYSMCTTAMFCIPLVL 232  
227 LLGHFGVLTGDEPQAGEPHEHQHSP 256  
Db 233 ILGCYGLI-----VRALIYKLDLNSP 253

RESULT 10  
US-10-092-135-6  
; Sequence 6, Application US/10092135  
; Publication No. US20030054374A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,  
; FILE REFERENCE: D0134.NP  
; CURRENT APPLICATION NUMBER: US/10/092,135  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: US 60/273,808  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 60/278,983  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 6  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: HOMO SAPIENS  
US-10-092-135-6

Query Match 16.9%; Score 326; DB 9; Length 373;  
Best Local Similarity 29.9%; Pred. No. 2.3e-23;  
Matches 93; Conservative 51; Mismatches 123; Indels 44; Gaps 10;  
QY 53 YLSLAYSIIIFILGLPLNGIVLHWSGQTKRWSCATTYLVNLMVADLLYVL-LPFLIITYS 111  
Db 53 YLPNAVILYFIIGFLGNSVAINMFVFMKPSGISVYMFNLALADFLYVLTLPALIFYF 112  
QY 112 LDDRWPFGELLCKLVHFLFYINLYGSILLTICISVHQFLGVCHPLGSLPYRTRHAWLGT 171  
Db 113 NKTDWIFGDVMCKLQRFIFHVNLVGSILFLTCISAHRYSGVVPYPLKSLGRKKNAIYVS 172  
QY 172 STTWALVVLQLLPTLAFSHTDY-INQMTIYDWTSONFDRFLPAYGIVLTLSGF-----LS 226  
Db 173 VLWLVVVAISPIIFYSGTGRKKNKTVCTYDSTDEYLRSYFIYSMCTTAMFCIPLVL 232  
QY 227 LLGHFGVLTGDEPQAGEPHEHQHSP 256  
Db 233 ILGCYGLI-----VRALIYKLDLNSP 253  
QY 286 HLLSAFSGLPALD---GSCGGLQD-----MEASGECEQLPQSPVLSFKGKN 331  
Db 276 FHVAKTNLRLARLDFTQTPANCAFNDRVYATYQTRGLASLNSC-----VDPILYFLAGDT 330  
QY 332 -RVRLLOKLRQ 341  
Db 331 FRRRLSRATRK 341  
RESULT 11

US-10-092-135-5  
; Sequence 5, Application US/10092135  
; Publication No. US20030054374A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,  
; FILE REFERENCE: D0134.NP  
; CURRENT APPLICATION NUMBER: US/10/092,135  
; PRIOR FILING DATE: 2002-03-06  
; PRIOR APPLICATION NUMBER: US 60/273,808  
; PRIOR FILING DATE: 2001-03-07  
; PRIOR APPLICATION NUMBER: US 60/278,983  
; PRIOR FILING DATE: 2001-03-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 5  
; LENGTH: 373  
; TYPE: PRT  
; ORGANISM: BOS TAURUS  
US-10-092-135-5

Query Match 16.9%; Score 326.5; DB 9; Length 373;  
Best Local Similarity 36.7%; Pred. No. 3.2e-23;  
Matches 77; Conservative 35; Mismatches 83; Indels 15; Gaps 4;  
QY 53 YLSLAYSIIIFILGLPLNGIVLHWSGQTKRWSCATTYLVNLMVADLLYVL-LPFLIITYS 111  
Db 53 YLPNAVILYFIIGFLGNSVAINMFVFMKPSGISVYMFNLALADFLYVLTLPALIFYF 112  
QY 112 LDDRWPFGELLCKLVHFLFYINLYGSILLTICISVHQFLGVCHPLGSLPYRTRHAWLGT 171  
Db 113 NKTDWIFGDVMCKLQRFIFHVNLVGSILFLTCISAHRYSGVVPYPLKSLGRKKNAIYVS 172  
QY 172 STTWALVVLQLLPTLAFSHTDY-INQMTIYDWTSONFDRFLPAYGIVLTLSGF-----LS 226  
Db 173 VLWLVVVAISPIIFYSGTGRKKNKTVCTYDSTDEYLRSYFIYSMCTTAMFCIPLVL 232  
QY 227 LLGHFGVLTGDEPQAGEPHEHQHSP 256  
Db 233 ILGCYGLI-----VRALIYKLDLNSP 253

RESULT 12  
US-10-270-144-4  
; Sequence 4, Application US/10270144.  
; Publication No. US20030049790A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; FILE REFERENCE: CL000750CON  
; CURRENT APPLICATION NUMBER: US/10/270,144  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/205,196  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 299  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-270-144-4

Query Match 16.5%; Score 320; DB 9; Length 299;  
Best Local Similarity 38.3%; Pred. No. 1e-22;  
Matches 72; Conservative 33; Mismatches 77; Indels 6; Gaps 3;  
QY 53 YLSLAYSIIIFILGLPLNGIVLHWSGQTKRWSCATTYLVNLMVADLLYVL-LPFLIITYS 111  
Db 12 YLPNAVILYFIIGFLGNSVAINMFVFMKPSGISVYMFNLALADFLYVLTLPALIFYF 71

Qy 112 LDDRWPGFGLCKLVHFLFYINLYSGISILLTLCISVHOFGLVCHPLCSPYRTRRHANLGT 171  
Db 72 NKTDFGDMACKLQRFTHFNLYSGISILLTLCISAHRYSGVVYPLKSLGRLKKNAIYVS 131  
Qy 172 STTWALVVLQPLPLASHT-DYINGOMIWDYMTSQENFDRFLFAYGIVLTLSGF-----LS 226  
Db 132 VLVWLVVVAISPILFYSYGTGRKNTVTCYDTTNSDYLRSFYISMCTTVAMFCPLVL 191  
Qy 227 LLGHFGLV 234  
Db 192 ILGCYGLI 199

RESULT 13  
US-09-782-974C-86  
; Sequence 86, Application US/09782974C  
; Publication No. US20030082534A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogeli, Gabriel  
; APPLICANT: Lind, Peter  
; APPLICANT: Wood, Linda S.  
; APPLICANT: Parodi, Luis A.  
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor  
; FILE REFERENCE: 411USPHRM311  
; CURRENT APPLICATION NUMBER: US/09/782,974C  
; PRIOR FILING DATE: 2002-09-04  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR FILING DATE: 09/714,449  
; PRIOR FILING DATE: 2000-11-16  
; PRIOR FILING DATE: 60/198,568  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR FILING DATE: 60/166,071  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR FILING DATE: 60/166,678  
; PRIOR FILING DATE: 1999-11-19  
; PRIOR FILING DATE: 60/173,396  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR FILING DATE: 60/184,129  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR FILING DATE: 60/185,421  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR FILING DATE: 60/185,554  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR FILING DATE: 60/186,530  
; PRIOR FILING DATE: 2000-03-02  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 86  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
9-782-974C-86

Query Match 14.5%; Score 281.5; DB 9; Length 336;  
Best Local Similarity 33.3%; Pred. No. 6.1e-19;  
Matches 59; Conservative 37; Mismatches 78; Indels 3; Gaps 3;

Qy 50 KQVYLSLAYSITIFILGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLVYL-LPFLII 108  
Db 32 KMHYLPVIYGIIFLVGPGNAVISTYIFKMRPKWSSTIIMNLACTDLYLTSLPFLIH 91  
Qy 109 TYSLDRWPFGELCKLVHFLFYINLYSGISILLTLCISVHOFGLVCHPLCSPYRTRRHAW 168  
Db 92 YYASGENWIFGDFMCKFIRSFHFNLYSSILFLTCFSIFRYCVIIHPMCSFCSIHKTRCAV 151  
Qy 169 LGTSTTWALVVLQPLP-TLAFSHDYINGOMIWDYMTSQENFDRFLFAYGIVLTLSGF 224  
Db 152 VACAVVWIIISLVAVIPMTFLTITSTNRTN-RSACLDTSSDELNTIKWYNLITATSTF 207

RESULT 14

US-10-023-775B-2  
; Sequence 2, Application US/10023775B  
; Publication No. US20030022282A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Inc. (EP/GB only)  
; APPLICANT: Pfizer Inc. (US, JP, EB except GB)  
; APPLICANT: Fidoock, Mark David  
; TITLE OF INVENTION: No. US20030022282A1el Polypeptide  
; FILE REFERENCE: PC10959AGPR  
; CURRENT APPLICATION NUMBER: US/10/023,775B  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: GB 0030854.4  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: US 60/260,590  
; PRIOR FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: US 60/296,660  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: GB 0111031.1  
; PRIOR FILING DATE: 2001-05-04  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-023-775B-2

Query Match 14.4%; Score 278.5; DB 9; Length 337;  
Best Local Similarity 32.8%; Pred. No. 1.2e-18;  
Matches 58; Conservative 38; Mismatches 78; Indels 3; Gaps 3;

Qy 50 KQVYLSLAYSITIFILGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLVYL-LPFLII 108  
Db 32 KMHYLPVIYGIIFLVGPGNAVISTYIFKMRPKWSSTIIMNLACTDLYLTSLPFLIH 91  
Qy 109 TYSLDRWPFGELCKLVHFLFYINLYSGISILLTLCISVHOFGLVCHPLCSPYRTRRHAW 168  
Db 92 YYASGENWIFGDFMCKFIRSFHFNLYSSILFLTCFSIFRYCVIIHPMCSFCSIHKTRCAV 151  
Qy 169 LGTSTTWALVVLQPLP-TLAFSHDYINGOMIWDYMTSQENFDRFLFAYGIVLTLSGF 224  
Db 152 VACAVVWIIISLVAVIPMTFLTITSTNRTN-RSACLDTSSDELNTIKWYNLITATSTF 207

RESULT 15  
US-10-270-144-2  
; Sequence 2, Application US/10270144  
; Publication No. US20030049790A1  
; GENERAL INFORMATION:  
; APPLICANT: Wei, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000750CON  
; CURRENT APPLICATION NUMBER: US/10/270,144  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/205,196  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Human  
US-10-270-144-2

Query Match 14.4%; Score 278.5; DB 9; Length 337;  
Best Local Similarity 32.8%; Pred. No. 1.2e-18;  
Matches 58; Conservative 38; Mismatches 78; Indels 3; Gaps 3;

Qy 50 KQVYLSLAYSITIFILGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLVYL-LPFLII 108  
Db 32 KMHYLPVIYGIIFLVGPGNAVISTYIFKMRPKWSSTIIMNLACTDLYLTSLPFLIH 91







GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:48:35 ; Search time 22 Seconds  
(without alignments)  
1573.109 Million cell updates/sec

Title: US-10-023-586B-2

Perfect score: 1936

Sequence: 1 MSLILLPSRSGSRRGAL.....QNKLGHPAGKRKPCGLNRS 360

oring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	400	20.7	328	2 I55450	G protein-coupled
2	387.5	20.0	328	2 JC4800	P2Y6 receptor - hu
3	374.5	19.3	375	2 A54946	P-2U nucleotide re
4	371	19.2	373	2 A47556	ATP receptor P2u -
5	367	19.0	365	2 S68679	G protein-coupled
6	346	17.9	362	2 S33733	G protein-coupled
7	328	16.9	373	2 JC4737	G protein-coupled
8	326.5	16.9	373	2 JC4162	P2Y receptor - bov
9	282.5	14.6	308	2 I50241	G protein-coupled
10	274.5	14.2	344	2 T09508	Intron 17 purinerg
11	258.5	13.4	370	2 JC5549	heptahelical P2Y5-
12	258	13.3	353	2 JC2492	G protein-coupled
13	248.5	12.8	391	2 A41795	somatostatin recep
14	247.5	12.8	359	2 S15403	angiotensin II rec
15	246	12.7	399	2 I48705	proteinase activat
16	245.5	12.7	391	2 A39297	somatostatin recep
17	244.5	12.6	359	2 S44425	angiotensin II rec
18	244.5	12.6	391	2 C41795	somatostatin recep
19	242.5	12.5	397	2 S66518	proteinase-activat
20	242	12.5	359	2 JC1104	angiotensin II rec
21	241.5	12.5	420	2 I51667	thrombin receptor
22	239	12.3	359	2 I39418	angiotensin II rec
23	238	12.3	359	2 A48857	angiotensin II rec
24	237.5	12.3	359	2 A42656	angiotensin II rec
25	237	12.2	359	2 JH0621	angiotensin II rec
26	237	12.2	359	2 JC2134	angiotensin II rec
27	234.5	12.1	354	2 I53033	G protein-coupled
28	234.5	12.1	362	2 S68207	G protein-coupled
29	234.5	12.1	362	2 A57641	G protein-coupled

30	231.5	12.0	423	2 JC7677	allatostatin recep
31	230.5	11.9	359	2 JC1194	angiotensin II rec
32	230.5	11.9	362	2 B57641	G protein-coupled
33	230	11.9	352	1 S27357	complement C5a ana
34	229	11.8	360	2 JC4587	chemokine (C-C) re
35	227	11.7	371	2 JC5498	G protein-coupled
36	226.5	11.7	359	2 JQ1516	angiotensin II rec
37	226.5	11.7	425	2 A37912	thrombin receptor
38	225.5	11.6	372	2 S26667	G protein-coupled
39	224.5	11.6	380	2 I38435	angiotensin recep
40	223	11.5	365	2 S68208	G protein-coupled
41	221	11.4	371	2 JC5796	probable chemoattr
42	221	11.4	398	2 I56504	mu opioid receptor
43	220.5	11.4	351	2 B42009	FMLP-related recep
44	220	11.4	398	2 I56517	mu opioid receptor
45	220	11.4	398	2 A57510	mu opioid receptor

ALIGNMENTS

RESULT 1

I55450

G protein-coupled P2 receptor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 19-May-2000

C;Accession: I55450

R;Chang, K.; Hanaoka, K.; Kumada, M.; Takuwa, Y.

J. Biol. Chem. 270, 26152-26158, 1995

A;Title: Molecular cloning and functional analysis of a novel P2 nucleotide receptor.

A;Reference number: I55450; MUID:96064682; PMID:7592819

A;Accession: I55450

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-328 <RES>

A;Cross-references: GB:D63665; NID:g1066007; PIDN:BAA09816.1; PID:g1066008

C;Superfamily: ATP receptor P2u

C;Keywords: G protein-coupled receptor

Query Match 20.7% Score 400; DB 2; Length 328;  
Best Local Similarity 32.5%; Pred. No. 6.7e-28;  
Matches 108; Conservative 53; Mismatches 137; Indels 34; Gaps 9;

Qy	29	MEKVDNMTSQQGL-----CFSEKRYKQVLSLAYSIIFILGLPLNGTLVHWSWGOTKRWS	84
Db	1	MER-DNGTIOAPGLPTTCVYREDEKRLLLPPVSVLVGLPLNVCVIAQICASRRILT	59
Qy	85	CATTYVNLVMADELYVL-LPFLIITYSLDDRWPFGELCKLVHFLFYINLYSGSILLTTC	143
Db	60	RSAVVTLNALADLLYACSLPLLIYNYARGDHPWPGDLACRLVRFVLANLHGSILFLTC	119
Qy	144	ISVHQFLGVCVHPLCSLPYR-TRRHAWLGTSTTVALVVLQLLPTLAFSTHTDYINGOMIWD	202
Db	120	ISFOYLGICHPHAPWHRKGRRAAWVGVVWLVTVAOCPLTAVFAATQTRNRTVCYD	179
Qy	203	MTSQENFDRLEAYGIVLFTLSGFL-----SLLGHFVGLFTDGPQDQARGEHEDRHSQS	258
Db	180	LSPPILSITRYLPYGNALIVIGFLPFTALLACYCRMARLRCRDGCPAGVAGERSKAAR	239
Qy	259	VHPDHTQWVPLHPLFCALPYHSLLLPHLL-SAFSGPLPDGSGCGGLQDMEASGSQL	317
Db	240	M-----AVVAAVFEVIFLPHITKTAVLVRSTPGVSCPVLFTFAAAYKGR	287
Qy	318	PQPS-----PVLSP-----KGGKNVRLLQKL	339
Db	288	PFASANSVLDPILFYFTQOKFRROPHDLLQKL	319

RESULT 2

JC4800

P2Y6 receptor - human

C;Species: Homo sapiens (man)

C;Date: 15-Oct-1995 #sequence\_revision 16-Aug-1996 #text\_change 17-Nov-2000

C:Accession: JC4800; G02514  
R:Communi, D.; Parmentier, M.; Boeynaems, J.M.  
Biochem. Biophys. Res. Commun. 222, 303-308, 1996  
A:Title: Cloning, functional expression and tissue distribution of the human P2Y6 receptor  
A:Reference number: JC4800; MUID:96222498; PMID:8670200  
A:Accession: JC4800  
A:Molecule type: mRNA  
A:Residues: 1-328 <COM>  
A:Cross-references: EMBL:X97058  
A:Experimental source: placenta  
R:Hammet, F.; Southey, M.C.; Somers, G.R.; Hutchins, A.M.; Venter, D.J.  
submitted to the EMBL Data Library, March 1996  
A:Reference number: H01373  
A:Accession: G02514  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 'M', 4-328 <HAM>  
A:Cross-references: EMBL:U52464; NID:g1407632; PIDN:AAB03572.1; PID:g1407633  
C:Genetics:  
C:Superfamily: ATP receptor P2u  
C:Keywords: glycoprotein; placenta; receptor; transmembrane protein  
F:26-52/Domain: transmembrane #status predicted <TM1>  
F:63-86/Domain: transmembrane #status predicted <TM2>  
F:104-122/Domain: transmembrane #status predicted <TM3>  
F:133-167/Domain: transmembrane #status predicted <TM4>  
F:193-216/Domain: transmembrane #status predicted <TM5>  
F:241-264/Domain: transmembrane #status predicted <TM6>  
F:283-305/Domain: transmembrane #status predicted <TM7>  
F:5,173/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match 20.0%; Score 387.5; DB 2; Length 328;  
Best Local Similarity 32.7%; Pred. No. 8.7e-27;  
Matches 108; Conservative 42; Mismatches 143; Indels 37; Gaps 9;

QY 33 DNMTSQEQL-----CFSEKRYKQVLSLAYSIIFILGLPLNGTVLWHSWGQTKRWSCATT 88  
DB 4 DNGTGALGLPPTTCVIRENFQKLLPPYSAVLAAGLPCLNVCITQICTSRALRTAV 63

QY 89 YLVNLMVADLLYL-LPFLIITYSLDDRPFGELCKLVHFLFYINLYSGILLTLCISVH 147  
DB 64 YTLNLALADLLYACSLPLIYNYAQGDHPFGDFGACRLVRFYANLHGSILFTCISFQ 123

QY 148 OFLGCHPLCSLPYR-TRRHAWLGTSTTVALVQLLPTLAFSHDTYINGOMIWDMTSQ 206  
DB 124 RYLIGCHPLAPWHKGRRAAWLVCVAVMLAVTTCPLTAIFAATGIQRNRTVCYDLSP 183

QY 207 ENFDRLFAYGIVLTSGFL-----SLIGHGVLFDTGQEPDQARGEPHEDRQSPSOVHPD 262  
DB 184 ALATHYMPYGMALTVIGFLLPNALLACVCLLAC-----RLCRQDGPAPVQAQ 231

QY 263 HPTGVWPLHPLFCALPYHSLLLPHLL-SAFSGLPALDGSOGLODMEASGECEQLPQP- 320  
DB 232 ERRGAAARMVAVAAFAISFLPEHTTKAYLAVRSTPCVPCVCTV--LEAFAAAYKGTREF 289

QY 321 -----SPVLSF-----KGGKNRVLLOKL 339  
DB 290 ASANSVLDPILFYFTQKKFRRRPHELLQKL 319

RESULT 3  
A54946  
P-20 nucleotide receptor - human  
C:Species: Homo sapiens (man)  
C:Date: 11-Nov-1994 #sequence\_revision 11-Nov-1994 #text\_change 17-Mar-1999  
C:Accession: A54946  
R:Parrr, C.E.; Sullivan, D.M.; Paradiso, A.M.; Lazarowski, E.R.; Burch, L.H.; Olsen, J.C.  
Proc. Natl. Acad. Sci. U.S.A. 91, 3275-3279, 1994  
A:Title: Cloning and expression of a human P-20 nucleotide receptor, a target for cystic  
A:Reference number: A54946; MUID:94211846; PMID:8159738  
A:Accession: A54946  
A:Status: preliminary  
A:Molecule type: mRNA; protein

A:Residues: 1-375 <PAR>  
A:Cross-references: GB:U07225  
A:Note: parts of this sequence were confirmed by protein sequencing  
C:Genetics:  
A:Gene: GDB:P2RY2; HP2U; P2U  
A:Cross-references: GDB:362713; OMIM:600041  
A:Map position: 11q13.5-11q14.1  
C:Superfamily: ATP receptor P2u  
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 19.3%; Score 374.5; DB 2; Length 375;  
Best Local Similarity 32.7%; Pred. No. 1.5e-25;  
Matches 115; Conservative 43; Mismatches 123; Indels 71; Gaps 13;

QY 43 CFSEKRYKQVLSLAYSIIFILGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLYL 102  
DB 25 CRFNEDEKVVLLPVSQVGVVCLGLCLNAGLVIFLCRLKTNASTYMFHLAVSDALYAA 84

QY 103 -LPFLIITYSLDDRPFGELCKLVHFLFYINLYSGILLTLCISVHOFGLGVCHPLCSLPY 161  
DB 85 SLPLLVYVYARGDHPFSTVLCVKLVRFYTNLYCSILFTCISVHRCGLVLRPLSLRW 144

QY 162 RTRRHAWLGTSTTVALVQLLPTLAFSHDTYINGOMIWDMTSQENFDRLFAYGIVLT 221  
DB 145 GRARYARRVAGAVVWLVLACQAPVLYFVTTT-ARGPLTCHDTSAPELFSRFVAYSSVM-- 201

QY 222 SGFLSLGHFGVLF-----TDGQEPDQARGEPHEDRQSPSOVHP 261  
DB 202 ---LGLL---FAVPFAVLVCYVLMARRLLKPAYGTSG-----GLPRAKRSVRT-- 245

QY 262 DHPTGVWPLHPLF--CALPYH-----SLLPHHLLSAFSGPLPALDGSOGLODME 309  
DB 246 -----IAVVLAVFALCFPLPFHVTRTLVYSFRSLDLSCHTFLNAIN-----MAYKVTRLA 293

QY 310 ASGSEQLPQPSPLSEFKGKNRVLLOKLQKNGEHPA-GRKRCPCGLNRS 360  
DB 294 SANS-----LDPVLYFLAGQRLVRFAADAKP-PTGSPSPATPARRTLGURS 339

RESULT 4  
A47556  
ATP receptor P2u - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 24-Sep-1999  
C:Accession: A47556  
R:Lustig, K.D.; Shiau, A.K.; Brake, A.J.; Julius, D.  
Proc. Natl. Acad. Sci. U.S.A. 90, 5113-5117, 1993  
A:Title: Expression cloning of an ATP receptor from mouse neuroblastoma cells.  
A:Reference number: A47556; MUID:93281707; PMID:7685114  
A:Accession: A47556  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-373 <LUS>  
A:Cross-references: GB:L14751; NID:g309457; PIDN:AAA39871.1; PID:g309458  
C:Superfamily: ATP receptor P2u  
C:Keywords: transmembrane protein

Query Match 19.2%; Score 371; DB 2; Length 373;  
Best Local Similarity 31.6%; Pred. No. 3e-25;  
Matches 113; Conservative 42; Mismatches 115; Indels 88; Gaps 12;

QY 43 CFSEKRYKQVLSLAYSIIFILGLPLNGTVLWHSWGQTKRWSCATTYLVNLMVADLLYL 102  
DB 25 CRFNEDEKVVLLPVSQVGVVCLGLCLNAGLVIFLCRLKTNASTYMFHLAVSDALYAA 84

QY 103 -LPFLIITYSLDDRPFGELCKLVHFLFYINLYSGILLTLCISVHOFGLGVCHPLCSLPY 161  
DB 85 SLPLLVYVYARGDHPFSTVLCVKLVRFYTNLYCSILFTCISVHRCGLVLRPLSLRW 144

QY 162 RTRRHAWLGTSTTVALVQLLPTLAFSHDTYINGOMIWDMTSQENFDRLFAYGIVLT 221  
DB 145 GRARYARRVAAVWVWLVLACQAPVLYFVTTTSGRTTCHDTSARELFSHFVAYSSVM-- 202



[illegible]

N;Alternate names: G-protein coupled receptor  
C;Species: Homo sapiens (man)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 24-Nov-1999  
C;Accession: T09508  
R;Bohm, S.K.; Trumpff, A.; Khitin, L.M.; Kong, W.; Payan, D.G.; Bunnett, N.W.  
submitted to the EMBL data Library, April 1997  
A;Description: The human purinergic receptor P2Y5 is encoded in intron 17 of the retinol  
A;Reference number: Z16705  
C;Accession: T09508  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-344 <BOH>  
A;Cross-references: EMBL:AF000546; NID:g2232068; PID:g2232069  
C;Genetics:  
A;Map position: 13  
Superfamily: ATP receptor P2u  
Keywords: G protein-coupled receptor; transmembrane protein

Query Match 14.2%; Score 274.5; DB 2; Length 344;  
Best Local Similarity 33.8%; Pred. No. 1.1e-16;  
Matches 70; Conservative 39; Mismatches 77; Indels 21; Gaps 7;

QY 32 VDNTSQEQGLQFSEKQVLSLAYSIIFILGLPLNGTVLHWSGQTKRWSCATTYLV 91  
DB 2 VSNSSH-----CFYNDFKTYTLGCMFSMVFLGLVSNCAIYIFICVLKVRNETTYMI 57  
QY 92 NLMAVDLLVYL-LPFLITYSLDDRWPFGLLCKLVHFLVINYLGSIILLTICISVHOFL 150  
DB 58 NLMSDLLFVETLPRF-FYTRNWFPGDLCKISVWLYTYMYGSIILFELTICISVDREL 116  
QY 151 GVCHPLCSLPYTRRRHAWLTSTTVALVLOLPT--LAFSHTDYINGQIMWYDMTSQEN 208  
DB 117 AIIVPEKSKTLRTKRNKIVCTGVMLTVIGGSAPAVEVQSTHSGNNASEACE-----EN 171  
QY 209 FD-----RLFAYGIVLTLSGLSLLGHF 231  
DB 172 FPEATWYTLRSRVI-----FIEIVGFF 194

RESULT 11  
JC5549  
heptahelical P2Y5-like receptor - human  
C;Species: Homo sapiens (man)  
C;Date: 02-Sep-1997 #sequence\_revision 05-Sep-1997 #text\_change 24-Sep-1999  
C;Accession: JC5549  
R;Janssens, R.; Boeynaens, J.M.; Godart, M.; Communi, D.  
Biochem. Biophys. Res. Commun. 236, 106-112, 1997  
A;Title: Cloning of a human heptahelical receptor closely related to the P2Y5 receptor.  
A;Reference number: JC5549; MUID:97366605; PMID:9223435  
A;Accession: JC5549  
A;Molecule type: DNA  
A;Residues: 1-370 <JAN>  
A;Cross-references: DDBJ:AF005419; NID:g2240034; PIDN:AAB66372.1; PID:g2240035  
C;Superfamily: ATP receptor P2u

Query Match 13.4%; Score 258.5; DB 2; Length 370;  
Best Local Similarity 33.2%; Pred. No. 3.1e-15;  
Matches 69; Conservative 35; Mismatches 93; Indels 11; Gaps 4;

QY 28 DMKVDNMTSQEQL-----CFSEKQVLSLAYSIIFILGLPLNGTVLHWSNGQ 79  
DB 8 DFQFQDSNSSLRPLRGNATANNCTIVDDSEKYNLNGAVISVVFILGNTNSVLSLFVFCFR 67  
QY 80 TKRWSCATTYLVNLMAVDLLVYL-LPFLITYSLDDRWPFGLLCKLVHFLVINYLGSI 138  
DB 68 MKMSETAFTINLAVSDLLFVCTLPKIP-YFNFRHWPFGDTLCKISGTAFLTNITGSM 126  
QY 139 LLLTICISVHOFLGVCHPLCSLPYTRRRHAWLTSTTVALVLOLPTLAFSHTDYINGQM 198  
DB 127 LFLTICISVDRELAIIVYFRSRTIRTRNSAIVCAGVILVLSGGISASLFTTWNVNTAT 186  
QY 199 IWYDMTSQENFDR-LFAYGIVLTLSGFL 225

Db 187 TCFEGLSKRWKTYLTKITIFIEVVGEI 214

RESULT 12  
JC2492  
G protein-coupled receptor 1 - rat  
N;Alternate names: GPR-1  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 16-Mar-1995 #sequence\_revision 26-May-1995 #text\_change 08-Oct-1999  
C;Accession: JC2492  
R;Marchese, A.; Cheng, R.; Lee, M.C.; Porter, C.A.; Heiber, M.; Goodman, M.; George, B.  
Biochem. Biophys. Res. Commun. 205, 1952-1958, 1994  
A;Title: Mapping studies of two G protein-coupled receptor genes: An amino acid difference  
A;Reference number: JC2492; MUID:95110347; PMID:7811287  
A;Accession: JC2492  
A;Molecule type: mRNA  
A;Residues: 1-353 <MAR>  
A;Cross-references: GB:S74702; NID:g786483; PIDN:AAB32978.1; PID:g786484  
C;Superfamily: vertebrate rhodopsin  
C;Keywords: glycoprotein; lipoprotein; phosphoprotein; receptor; thiolester bond; transmembrane  
F;74-94/Domain: transmembrane #status predicted <TM2>  
F;112-133/Domain: transmembrane #status predicted <TM3>  
F;134-135/Region: DR motif  
F;155-175/Domain: transmembrane #status predicted <TM4>  
F;209-229/Domain: transmembrane #status predicted <TM5>  
F;246-266/Domain: transmembrane #status predicted <TM6>  
F;295-306/Domain: transmembrane #status predicted <TM7>  
F;14-273/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;150,231/Binding site: phosphate (Thr) (covalent) (by protein kinase A) #status predicted  
F;330/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 13.3%; Score 258; DB 2; Length 353;  
Best Local Similarity 30.8%; Pred. No. 3.2e-15;  
Matches 72; Conservative 41; Mismatches 81; Indels 40; Gaps 8;

QY 26 SRDM--EKVDNMTSQEQLQFSEKQVLS-----LAYSIIFILGLPLNGTVLW-- 74  
DB 4 SREMLFEELDNYSYALEYYSQEPDAEENYVPGIVHVSLLLYALAFVLGIPGNAIVWF 63  
QY 75 -HSGQTKRWSCATTYLVNLMAVDLLVYL-LPFLITYSLDDRWPFGLLCKLVHFLFYI 132  
DB 64 GFKWKKT----VTTLWFENLAIADIFIVFLPLYSIVVALSFHWPFGRLCKLSFIAQL 119  
QY 133 NLXGSIILLTICISVHOFLGVCHPLCSLPYTRRRHAWLTSTTVALVLOLPTLAFSHTD 192  
DB 120 NMESVFFELVISLDRVYIHLIHPGLSHPHRTLKNSLLVFLVWLLASLLGGPTLYFRDTV 179  
QY 193 YINGQIMWYDMTSQENFDRL-----FAYGIVLT-----SGFLSLL 228  
DB 180 EVNNRIICYN-----NFQYEVLTLMRHVLTTWVKFLFGYLLPLLTMSSCYLCI 228

RESULT 13  
A41795  
somatostatin receptor 1 - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 24-Nov-1999  
C;Accession: A41795  
R;Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Seino, S.  
Proc. Natl. Acad. Sci. U.S.A. 89, 251-255, 1992  
A;Title: Cloning and functional characterization of a family of human and mouse somatostatin  
A;Reference number: A41795; MUID:92108031; PMID:1346068  
A;Accession: A41795  
A;Molecule type: DNA  
A;Residues: 1-391 <YAM>  
A;Cross-references: GB:M81829; NID:g307433; PIDN:AAA58247.1; PID:g307434  
A;Note: sequence extracted from NCBI backbone (NCBIN:74767, NCBI:P:74768)  
C;Genetics:  
A;Gene: GDB:SSTR1  
A;Cross-references: GDB:134185; OMIM:182451  
A;Map position: 14q13-14q13  
A;Introns: #status absent  
C;Superfamily: vertebrate rhodopsin

C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; phd  
F:58-84/Domain: transmembrane #status predicted <TM1>  
F:95-120/Domain: transmembrane #status predicted <TM2>  
F:132-153/Domain: transmembrane #status predicted <TM3>  
F:173-195/Domain: transmembrane #status predicted <TM4>  
F:220-230/Domain: transmembrane #status predicted <TM5>  
F:269-296/Domain: transmembrane #status predicted <TM6>  
F:302-326/Domain: transmembrane #status predicted <TM7>  
F:4,44,48,381/Binding site: carboxylate (Asn) (covalent) #status predicted  
F:130-208/bisulfide bonds: #status predicted  
F:172/Binding site: phosphate (Thr) (covalent) (by CAMP-dependent kinase) #status predicted  
F:265/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status predicted  
F:339/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 12.8%; Score 248.5; DB 2; Length 391;  
Best Local Similarity 32.6%; Pred. No. 2.5e-14;  
Matches 74; Conservative 41; Mismatches 103; Indels 9; Gaps 6;

Qy 7 PSRGS--RSGSRGALLLEGASRDMEKVDMMNTSQOGLCFSEKQYVLSIAYSIIFIL 64  
Db 16 PSPGCGEGGSGRGP--GAGAAGDMEEPGRNASQNGTLSE--GQGSAILISFIYSVWCLV 71

Qy 65 GLPLNGTVLHWSWGOTKRWSCATTVLVNLMVAD-LLYVLLPFLIITYSLDDRWPFCELLC 123  
Db 72 GLCGNSMVIYVILRYAKRKATNIYILNLAIADELLMLSVFPL-VTSTLLRHWPFGALLC 130

Qy 124 KLVHFLFVINYLSILLTLCISVHQFLGVCPLCSLPYTRRRHAWLGSTTWALVVLQLL 183  
Db 131 RLVLSDVADVNMTSIYCLTVLSVDYVAVVHPPIKARVRRPTVAKVNVGVVLSLVIL 190

Qy 184 PTLAFSHTDYINGQMIWYDMTSQENFDR-LFAYGIVLTLSGFLSLG 229  
Db 191 PTVPFSRTAANSDDGTVAACNMLMPEAQRWLVGFVLYTFLMGFLLPVG 237

RESULT 14  
S15403  
angiotensin II receptor type 1 - bovine  
C:Species: Bos primigenius indicus x Bos primigenius taurus (cattle)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 24-Nov-1999  
C:Accession: S15403  
R:Sasaki, K.; Yamano, Y.; Bardhan, S.; Iwai, N.; Murray, J.J.; Hasegawa, M.; Matsuda, Y.  
Nature 351, 230-233, 1991  
A:Title: Cloning and expression of a complementary DNA encoding a bovine adrenal angiotensin II receptor type 1  
A:Reference number: S15403; MUID:91251900; PMID:2041569  
A:Accession: S15403  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-359 <SAS>  
A:Cross-references: GB:X62294; NID:g43; PIDN:CAA44182.1; PID:g44  
C:Superfamily: vertebrate rhodopsin

Query Match 12.8%; Score 247.5; DB 2; Length 359;  
Best Local Similarity 31.0%; Pred. No. 2.8e-14;  
Matches 63; Conservative 42; Mismatches 87; Indels 11; Gaps 4;

Qy 34 MNTSQOGL-----CFSEKQYVLSL--AYSIIIFILGPLNGTVLHWSWGOTKRWSC 85  
Db 3 LNSSTDGIGKRTDDCPKAGRHNHYIFIMPTLYSIIFVVGIFGNSLVIVIFYMVKLTV 62

Qy 86 ATTYLVNLMVADLLYL-LPFLIITYSLDDRWPFCELLKLVHFLFYINLYGSILLTLCI 144  
Db 63 ASVFLLNLALADLCFLLLPLWAVVATMEYRWPFGNYLCKIASASVSFNLYASVFLTCL 122

Qy 145 SVHOFGLGVCPLCSLPYTRRRHAWLGSTTWALVVLQLLPTLAFSHTDYINGQMIWYDMT 204  
Db 123 SIDRYLAIVHPKSLRRTMLVAKVTCIIIIWLLAGLASLPTIIHNRVFFIENTNITVCAF 182

Qy 205 SQENFDRLFAYGIVLT--LSGFL 225  
Db 183 HYESQNSTLPVGLGTLTKNLGFL 205

RESULT 15  
I48705  
proteinase activated receptor 2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 24-Nov-1999  
C:Accession: I48705  
R:Nystedt, S.; Larsson, A.K.; Aberg, H.; Sundelin, J.  
J. Biol. Chem. 270, 5950-5955, 1995  
A:Title: The mouse proteinase-activated receptor-2 cDNA and gene. Molecular cloning a  
A:Reference number: I48705; MUID:95197620; PMID:7890726  
A:Accession: I48705  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-399 <RES>  
A:Cross-references: EMBL:Z48043; NID:9663020; PIDN:CAA88097.1; PID:9663021  
C:Superfamily: ATP receptor P2u

Query Match 12.7%; Score 246; DB 2; Length 399;  
Best Local Similarity 35.1%; Pred. No. 4.3e-14;  
Matches 68; Conservative 34; Mismatches 80; Indels 12; Gaps 7;

Qy 48 KYQVYLSIAYSIIFILGPLNGTVLHWSWGOTKRWSCATTVLVNLMVADLLYLPLFLI 107  
Db 74 KLTTFELPVVIVIVIGLPSNGMALIFLFRKKHFAVIYMANLADLLSVIWFPLK 133

Qy 108 ITYSL-DDRWPFCELLKLVHFLFYINLYGSILLTLCISVHQFLGVCPLCSLPYTRRH 166  
Db 134 ISYHLHGNWVYGEALCKVLIGFFYGNMYCSILFMTCLSVQRYWIVNPM-GHP-RKKN 191

Qy 167 AWLGTS-TTWALVVLQLLPTLAFSHTDYINGQMI--WYDMTSQENF-DRLFAYGIVLTL 222  
Db 192 IAVGVSLAIWLLIFLVTIPLYVMKOTIYPALNITTCCHDVLPEEVLVGDMMFNFLSLAIG 251

Qy 223 GFLSLLGHEGVLEF 236  
Db 252 VFL-----FPALLT 260

Search completed: May 22, 2003, 12:51:13  
Job time : 23 secs

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 22, 2003, 12:47:10 ; Search time 14 Seconds  
(without alignments)  
1066.534 Million cell updates/sec

Title: US-10-023-586B-2

Perfect score: 1936

Sequence: 1 MSLILPSRSGSRGAL.....QNKLGHPAGRKRCPLNRS 360

oring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	447	23.1	328	1 P2Y3 CHICK	Q98907 gallus gall
2	447	23.1	328	1 P2Y3 MELGA	Q93361 melagris g
3	416	21.5	537	1 P2Y8 XENLA	P79928 xenopus lae
4	400	20.7	328	1 P2Y6 RAT	O63371 rattus norv
5	388	20.0	361	1 P2Y4 RAT	O35811 rattus norv
6	387.5	20.0	328	1 P2Y6 HUMAN	Q15077 homo sapien
7	387.5	20.0	377	1 P2Y2 HUMAN	P41231 homo sapien
8	386	19.9	361	1 P2Y4 MOUSE	Q9JJ57 mus musculu
9	371	19.2	373	1 P2Y2 MOUSE	P35383 mus musculu
10	367	19.0	365	1 P2Y4 HUMAN	P51582 homo sapien
11	366.5	18.9	374	1 P2Y2 RAT	P41232 rattus norv
12	346	17.9	362	1 P2YR CHICK	P34996 gallus gall
13	346	17.9	362	1 P2YR MELGA	P49652 melagris g
14	329.5	17.0	373	1 P2YR RAT	P49651 rattus norv
15	328	16.9	373	1 P2YR HUMAN	P47900 homo sapien
16	326.5	16.9	373	1 P2YR BOVIN	P48042 bos taurus
17	320	16.5	373	1 P2YR MOUSE	P49650 mus musculu
18	282.5	14.6	308	1 P2Y5 CHICK	P32250 gallus gall
19	274.5	14.2	344	1 P2Y5 HUMAN	P43657 homo sapien
20	260.5	13.5	165	1 P2Y4 CRIGR	P58826 cricetus
21	258.5	13.4	370	1 P2Y9 HUMAN	Q99677 homo sapien
22	258	13.3	353	1 GPRI RAT	P46090 rattus norv
23	252	13.0	353	1 APJ XENLA	P79960 xenopus lae
24	250.5	12.9	367	1 GP17 HUMAN	Q13304 homo sapien
25	250.5	12.9	374	1 PAR3 HUMAN	O00254 homo sapien
26	249	12.9	397	1 PAR2 RAT	Q63645 rattus norv
27	248.5	12.8	359	1 AG2R SHEEP	O77590 ovis aries
28	248.5	12.8	391	1 SSRI HUMAN	P30872 homo sapien
29	247.5	12.8	359	1 AG2R BOVIN	P25104 bos taurus
30	246	12.7	399	1 PAR2 MOUSE	P55086 mus musculu
31	245.5	12.7	391	1 SSRI MOUSE	P28846 rattus norv
32	244.5	12.6	359	1 AG2R CANFA	P43240 canis famil
33	244.5	12.6	391	1 SSRI MOUSE	P30873 mus musculu

RESULT 1				
P2Y3_CHICK	34	242.5	12.5	397
ID	35	242	12.5	359
AC	36	241.5	12.5	420
Q98907;	37	241	12.4	359
DT	38	239	12.3	359
01-NOV-1997 (Rel. 35, Created)	39	238	12.3	359
01-NOV-1997 (Rel. 35, Last sequence update)	40	237.5	12.3	359
16-OCT-2001 (Rel. 40, Last annotation update)	41	237	12.2	359
P2Y purinoceptor 3 (P2Y3) (Nucleoside diphosphate receptor).	42	237	12.2	359
P2Y3.	43	237	12.2	359
GN	44	236.5	12.2	377
Gallus gallus (Chicken).	45	236	12.2	374
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Archosauria; Aves; Neognathae; Galliformes; Phasianinae;				
Gallus.				
NCBI_TaxID=9031;				
[1]				
SEQUENCE FROM N.A.				
TISSUE=Brain;				
MEDLINE=96319774; PubMed=8700132;				
Webb T.E., Henderson D., King B.F., Wang S., Simon J.,				
Bateson A.N., Burnstock G., Barnard E.A.;				
"A novel G protein-coupled P2 purinoceptor (P2Y3) activated				
preferentially by nucleoside diphosphates.";				
Mol. Pharmacol. 50:258-265(1996).				
-1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADP > UTP > ATP - UDP. THE				
ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE				
A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.				
-1- SUBCELLULAR LOCATION: Integral membrane protein.				
-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.				
-----				
This SWISS-PROT entry is copyright. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL outstation -				
the European Bioinformatics Institute. There are no restrictions on its				
use by non-profit institutions as long as its content is in no way				
modified and this statement is not removed. Usage by and for commercial				
entities requires a license agreement (See http://www.isb-sib.ch/announce/				
or send an email to license@isb-sib.ch).				
-----				
EMBL; X98283; CAA66930.1; -.				
HSSP; P34996; 1DDO.				
InterPro: IPR000276; GPCR_Rhodpsn.				
Pfam: PF00001; 7tm1.1; 1.				
PRINTS; PR00237; GPCRHOOPS.				
PROSITE; PS00237; G_PROTEIN_RECP_FL_1; FALSE_NEG.				
PROSITE; PS0262; G_PROTEIN_RECP_FL_2; 1.				
G-protein coupled receptor; Transmembrane; Glycoprotein.				
DOMAIN 1 22 EXTRACELLULAR (POTENTIAL).				
TRANSMEM 23 43 1 (POTENTIAL).				
DOMAIN 44 57 CYTOPLASMIC (POTENTIAL).				
TRANSMEM 58 78 2 (POTENTIAL).				
DOMAIN 79 96 EXTRACELLULAR (POTENTIAL).				
TRANSMEM 97 117 3 (POTENTIAL).				
DOMAIN 118 139 CYTOPLASMIC (POTENTIAL).				
TRANSMEM 140 160 4 (POTENTIAL).				
DOMAIN 161 189 EXTRACELLULAR (POTENTIAL).				
TRANSMEM 190 210 5 (POTENTIAL).				
DOMAIN 211 231 CYTOPLASMIC (POTENTIAL).				

```
FT TRANSMEM 232 252 6 (POTENTIAL).
FT DOMAIN 253 275 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 276 298 7 (POTENTIAL).
FT DOMAIN 299 323 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 94 172 BY SIMILARITY.
SQ SEQUENCE 328 AA; 37586 MW; 7A3BFIC91F54FAAB CRC64;

Query Match 23.1%; Score 447; DB 1; Length 328;
Best Local Similarity 45.9%; Pred. No. 2.1e-22;
Matches 85; Conservative 31; Mismatches 67; Indels 2; Gaps 2;

QY 43 CQFSKYKQVLSLAYSIIFILGLPLNGTVLWHSQTKRWSCATTYLVNLMVADLLYL 102
DB 13 CTFHEEFKQVLLPLVYSVVFVLLGLPLNAVIGQIWLARKALTRTTIYMLNADLLYVC 72
QY 103 -LPFLIITYSLDDRPFCELLCKLVHFLFYINLYGSIILLTICISVHQFLGVCPLCSL-P 160
DB 73 SLPLLIYNTQKDYWPFQDFCTCKFRFYTNLHGSILFLTICISVORYMGICHPLASWHK 132
QY 161 YTRRHAWLGTSTTVALVQLLPTLAFSHDTYINGOMIWMYDMSQENFDRFLFAYGIVLT 220
DB 133 KKGKLLTLVCAAVWFIVIAQCLPTFVFASTGTQRNRTVCYDLSPPDRSTSYFPYGITLT 192
221 LSGFL 225
193 IIGFL 197

RESULT 2
P2Y3_MEIGA
ID P2Y3_MEIGA STANDARD; PRT; 328 AA.
AC O93361;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE P2Y purinoceptor 3 (P2Y3) (Nucleoside diphosphate receptor).
GN P2RY3.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98401046; PubMed=9730913;
RA Li Q., Olesky M., Palmer R.K., Harden T.K., Nicholas R.A.;
RT "Evidence that the p2y3 receptor is the avian homologue of the
mammalian p2y6 receptor.";
RL Mol. Pharmacol. 54:541-546(1998).
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR UDP< ADP = UTP. THE ACTIVITY
OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A
PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; AF069555; AAC23863.1; -
DR HSSP; P34996; 1DDD.
DR InterPro; IPR00276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G-PROTEIN_RECPEP_F1_1; FALSE_NEG.
DR PROSITE; PS00262; G-PROTEIN_RECPEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 22 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 23 43 1 (POTENTIAL).
```

```
FT DOMAIN 44 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 2 (POTENTIAL).
FT DOMAIN 79 96 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 97 117 3 (POTENTIAL).
FT DOMAIN 118 139 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 140 160 4 (POTENTIAL).
FT DOMAIN 161 189 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 190 210 5 (POTENTIAL).
FT DOMAIN 211 231 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 232 252 6 (POTENTIAL).
FT DOMAIN 253 275 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 276 298 7 (POTENTIAL).
FT DOMAIN 299 323 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 94 172 BY SIMILARITY.
SQ SEQUENCE 328 AA; 37594 MW; B74D49B99C7164A5 CRC64;

Query Match 23.1%; Score 447; DB 1; Length 328;
Best Local Similarity 44.6%; Pred. No. 2.1e-22;
Matches 86; Conservative 32; Mismatches 73; Indels 2; Gaps 2;

QY 35 NTSBOGLCQFSEKYKQVLSLAYSIIFILGLPLNGTVLWHSQTKRWSCATTYLVNLM 94
DB 5 NTAGRNCTQEEFKQVLLPLVYSVVFLLGLPLNAVIGQIWLARKALTRTTIYMLNLA 54
QY 95 VADLLYL-LPFLIITYSLDDRPFCELLCKLVHFLFYINLYGSIILLTICISVHQFLGVC 153
DB 65 TADLLYVCSLPLLIYNTQKDYWPFQDFCTCKFRFYTNLHGSILFLTICISVORYMGIC 124
QY 154 HPLCSL-PYTRRHAWLGTSTTVALVQLLPTLAFSHDTYINGOMIWMYDMSQENFDR 212
DB 125 HPLASWHKKGKLLTLVCAAVWFIVIAQCLPTFVFASTGTQRNRTVCYDLSPPDRSASY 184
QY 213 FAYGIVLTLSGFL 225
DB 185 FPGITLTITGFL 197

RESULT 3
P2Y8_XENLA
ID P2Y8_XENLA STANDARD; PRT; 537 AA.
AC P79928;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE P2Y purinoceptor 8 (P2Y8).
GN P2RY8.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Neural plate;
RX MEDLINE=97284734; PubMed=9139711;
RA Bogdanov Y.D., Dale L., King B.F., Whittock N., Burnstock G.;
RT "Early expression of a novel nucleotide receptor in the neural plate
of Xenopus embryos.";
RL J. Biol. Chem. 272:12583-12590(1997).
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ATP, UTP, GTP AND ITP.
ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
```



```

or send an email to license@lsb-sib.ch).
-----
CC EMBL; X99953; CAA68213.1; -.
CC DR HSSP; P34996; 1DDD.
CC DR InterPro; IPR000276; GPCR_Rhodopsn.
CC DR Pfam; PF00001; 7tm_1; 1.
CC DR PRINTS; PR00237; GPCRHOPOPSN.
CC DR PROSITE; PS00237; G-PROTEIN_REC_F1_1; 1.
CC DR PROSITE; PS00262; G-PROTEIN_REC_F2_1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT DOMAIN 1 49 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 50 70 1 (POTENTIAL).
CC FT DOMAIN 71 79 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 80 100 2 (POTENTIAL).
CC FT DOMAIN 101 118 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 119 139 3 (POTENTIAL).
CC FT DOMAIN 140 161 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 162 182 4 (POTENTIAL).
CC FT DOMAIN 183 210 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 211 231 5 (POTENTIAL).
CC FT DOMAIN 232 254 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 255 275 6 (POTENTIAL).
CC FT DOMAIN 276 292 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 293 316 7 (POTENTIAL).
CC FT DOMAIN 317 337 CYTOPLASMIC (POTENTIAL).
CC FT DISULFID 116 193 BY SIMILARITY.
CC FT CARBOHYD 26 26 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 29 29 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 537 AA; 62024 MW; B2CF24812F3C19F2 CRC64;

Query Match 21.5%; Score 416; DB 1; Length 537;
Best Local Similarity 31.8%; Pred. No. 3.3e-20;
Matches 110; Conservative 59; Mismatches 139; Indels 38; Gaps 10;

QY 31 KVDNN-TSQGGLCORSEKVKQVLSLAYSIIFILGLPLNGTVLWHSWGOTKRWSCATTY 89
DB 22 KLLMNLNTDIEDICVDEGEKFLPLPSYSAVFWGLPLNIAAWIFIAKMRPNPTTVY 81

QY 90 LVNLWVADLLYVL-LPELIITYSLDDRWPEGLCKLHFLFYINLYGSIILLTICISVHQ 148
DB 82 MFNLASDITLVLSPLTVVYVADKNNWPEGLCKLVRELFYANLYSSILFTICISVHR 141

QY 149 FLGVCHPLCSLPYTRRHANLGTSTWALVVLQPLTAFSHDYINGQMIWDMTSQEN 208
DB 142 YRGVCHPITSRLRNNAKHAYVICALVWLSVTCLVPLNIEFTVSPKYNKNTCHDTRPED 201

QY 209 FDRLEFAY--GIVLTLSGF--LSLGHFGSVLFTDQGEPDQARGEPEHDSQVHPDHP 264
DB 202 FARVEYSTAIMCLLFGIPCLTIAGCYGLMTRMLMKDIVS-----GNOQTLPYKKRSIK 256

QY 265 TGVWPLHPLFCALPYHSLLLPHLLSAFSGPLALDGSQC-----GLQDMEASGE 313
DB 257 T-----IIFVMAFAICFEPFHITRTLYYYARLLGKICVALNINVTYKVRPLASANS 310

QY 314 CEQLPQSPVLSF-KGKNRVRLLQKLRQKNLGEHPAGRKRCPLGN 358
DB 311 C-----IDPILYFLANDRYRRRLIRTVRRRS-----SVENRRRCMHTN 347

RESULT 4
P2Y6_RAT STANDARD; PRT; 328 AA.
AC Q63371;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE P2Y purinocceptor 6 (P2Y6).
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Aortic smooth muscle;
RX MEDLINE=96064682; PubMed=7592819;
RA Chang K., Hanaoka K., Kumada M., Takawa Y.;
RT "Molecular cloning and functional analysis of a novel P2 nucleotide
RT receptor.";
RL J. Biol. Chem. 270:26152-26158(1995).
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR UTP > ADP = 2-METHYLTHIO-ATP
CC > ADP-BETA-S > ATP = ATP-GAMMA-S. THE ACTIVITY OF THIS RECEPTOR IS
CC MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM. FUNCTIONALLY COUPLED TO
CC PHOSPHOLIPASE C.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN VARIOUS TISSUES
CC INCLUDING LUNG, STOMACH, INTESTINE, SPLEEN, MESENTERY, HEART, AND,
CC MOST PROMINENTLY, AORTA.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC -----
CC EMBL; D63665; BAA09816.1; -.
CC DR HSSP; P34996; 1DDD.
CC DR InterPro; IPR000276; GPCR_Rhodopsn.
CC DR Pfam; PF00001; 7tm_1; 1.
CC DR PRINTS; PR00237; GPCRHOPOPSN.
CC DR PROSITE; PS00237; G-PROTEIN_REC_F1_1; FALSE_NEG.
CC DR PROSITE; PS00262; G-PROTEIN_REC_F2_1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein.
CC FT DOMAIN 1 27 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 28 48 1 (POTENTIAL).
CC FT DOMAIN 49 62 2 (POTENTIAL).
CC FT TRANSMEM 63 83 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 84 101 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 102 122 3 (POTENTIAL).
CC FT DOMAIN 123 144 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 145 165 4 (POTENTIAL).
CC FT DOMAIN 166 194 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 195 215 5 (POTENTIAL).
CC FT DOMAIN 216 236 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 237 257 6 (POTENTIAL).
CC FT DOMAIN 258 280 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 281 303 7 (POTENTIAL).
CC FT DOMAIN 304 328 EXTRACELLULAR (POTENTIAL).
CC FT DISULFID 99 177 BY SIMILARITY.
CC FT CARBOHYD 5 5 N-LINKED (GLCNAC...) (POTENTIAL).
CC SQ SEQUENCE 328 AA; 36677 MW; 58EBE1997038BA5B CRC64;

Query Match 20.7%; Score 400; DB 1; Length 328;
Best Local Similarity 32.5%; Pred. No. 2.3e-19;
Matches 108; Conservative 53; Mismatches 137; Indels 34; Gaps 9;

QY 29 MEKVDMTSQEGL----COFSEKVKQVLSLAYSIIFILGLPLNGTVLWHSWGOTKRWS 84
DB 1 MER-DNGTIOAPGLPTTCVIREDFKRLLPVYVGLVGLPLNVCVIAQICASRRFLT 59

QY 85 CATTYLVNLVADLLYVL-LPFLIITYSLDDRWPEGLCKLHFLFYINLYGSIILLTIC 143
DB 60 RSAVYTLNLALADLLYACSLPLLIYNYARGDHPFGDLACRLVRELFYANLHGSILFTC 119

QY 144 ISVHQFLGVCHPLCSLPYR-TRRHANLGTSTWALVVLQPLTAFSHDYINGQMIWYD 202
DB 120 ISFQYLGICHPLAPWHKRGRRAAVWVCGVWLVVTAQCLPTAFVAAFGIQRNRTCYD 179

QY 203 MTSQENEDRLFAVGIVLTLSGFL----SLIGHFGVLTGQGEPDQARGEPEHDSQSPSQ 258
DB 180 LSPPISTRYLPYGMALTIVGFLPFTALLACACYCRMARRLCRDGPAGPVAQERRSKAR 239

```

QY 259 VHPDPTGWPVLPFCALPYHSLLLPHLL-SAFSGLPALDGSQCGLODMASGECESOL 317  
 Db 240 M-----AVVAAVFISELPFHITKTAYLAVRSTPGVSCPVLTFFAAAKGTGTR 287  
 QY 318 PQPS-----PVLSF-----KGGKNRVRLLOKL 339  
 Db 288 PPASANSVLDPIFYFTQOKFRQRPHDLLQKL 319

RESULT 5  
 P2Y4\_RAT  
 ID P2Y4\_RAT STANDARD; PRT; 361 AA.  
 AC O35811;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE P2Y purinoceptor 4 (P2Y4).  
 GN P2Y4 OR P2Y4  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP STRAIN-Sprague-Dawley; TISSUE=Liver;  
 Bogdanov F.D., Wildman S., King B.F., Burnstock G.;  
 Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-Sprague-Dawley; TISSUE=Brain;  
 RX MEDLINE=98421785; PubMed=9751165;  
 RA Webb T.E., Henderson D., Roberts J.A., Barnard E.A.;  
 RT "Molecular cloning and characterization of the rat P2Y4 receptor.";  
 RL J. Neurochem. 71:1424-1434(1998).  
 CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that  
 activate a phosphatidylinositol-calcium second messenger system.  
 CC Not activated by ADP or UDP.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Widely expressed at low levels. In brain,  
 CC higher expression in the pineal gland and ventricular system.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Y14705; CAA75007.1; -;  
 DR EMBL; Y11433; CAA72241.1; -;  
 DR HSSP; P34996; 1DDD.  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm1.1;  
 DR PRINTS; PS00237; GPCRRHODPSN  
 DR PROSITE; PS00237; G-PROTEIN\_RECEPTOR\_F1\_1; 1.  
 DR PROSITE; PS0262; G-PROTEIN\_RECEPTOR\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 30  
 FT TRANSMEM 31 58  
 FT DOMAIN 59 68  
 FT TRANSMEM 69 91  
 FT DOMAIN 92 108  
 FT TRANSMEM 109 127  
 FT DOMAIN 128 149  
 FT TRANSMEM 150 170  
 FT DOMAIN 171 192  
 FT TRANSMEM 193 218  
 FT DOMAIN 219 242  
 FT TRANSMEM 243 265  
 FT DOMAIN 266 283  
 FT TRANSMEM 284 305

FT DOMAIN 306 361 CYTOPLASMIC (POTENTIAL).  
 FT DISULFID 104 181 BY SIMILARITY.  
 FT CARBOHYD 175 175 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 361 AA; 40893 MW; 0377F96E54B449A3 CRC64;

Query Match 20.0%; Score 388; DB 1; Length 361;  
 Best Local Similarity 32.1%; Pred. No. 1.5e-18;  
 Matches 105; Conservative 54; Mismatches 126; Indels 42; Gaps 9;

QY 37 SOEOGLCOFSEKYKOVYLSLAYSIIFILGPLNGVTLVHWSGQTKRWSGATYLYNLMAVA 96  
 Db 17 SSGDGDRENEEFKILLPMSTAVVVFVGLALNAPTLMFLFRLRPWDATATYMFHLALS 76  
 QY 97 DLLYVL-LPFLIITYSLDDRMPPFGELCKLVHFLFYINLYSGISILLTLCISVHQFVGCHP 155  
 Db 77 DTLYVLSLPTLVVYVAARNHWPFGTGLCKFVRFYWNLYCSVLFTCTSVHRYLGICH 136  
 QY 156 LCSLPYRTRRHAWLGTSTWALVQLLPTLAFSTHTDINGOMIWDYMTSQENFDR--LF 213  
 Db 137 LRATRWGRPRFASLLCLGLVWVAGCLVFNLFVFTTNANGFTILCHDTLPPEFDHYVYF 196  
 QY 214 YGIVLTLSG--FLSLGLHFGVLFDTGQEPDQARCEPHEDRQHSQVHPDHTGWPPLH 271  
 Db 197 SSANVLLFGLPFLITLVYCYGLMARRLYRPLPGAQO-----SSRLSLRTIAVLT 249  
 QY 272 PLFCALPYH---SLLPHHLLSAFSGLPALDGSQCGL-----ODMEASGECESOL 317  
 Db 250 FAVCFVPEPHITRTIYYQARLLQA-----DCHVLNVNVVYKVRPLASANS--- 296

QY 318 PQSPVL-SFKGKNRVRLLQKLQKRONK 343  
 Db 297 --LDPVLYFTGDKYRNQLQQLCRGSK 321

RESULT 6  
 P2Y6\_HUMAN  
 ID P2Y6\_HUMAN STANDARD; PRT; 328 AA.  
 AC Q15077; Q15754;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE P2Y purinoceptor 6 (P2Y6).  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Communi D., Parmentier M., Boeynaems J.M.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97432828; PubMed=9286708;  
 RA Somers G.R., Hammet F., Woollatt E., Richards R.I., Southey M.C.,  
 RA Venter D.J.;  
 RT "Chromosomal localization of the human P2Y6 purinoceptor gene and  
 RT phylogenetic analysis of the P2Y purinoceptor family.";  
 RL Genomics 44:127-130(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=Placenta;  
 RX MEDLINE=97415792; PubMed=9268704;  
 RA Maier R., Glatz A., Mosbacher J., Bilbe G.;  
 RT "Cloning of P2Y6 cDNAs and identification of a pseudogene: comparison  
 RT of P2Y receptor subtype expression in bone and brain tissues.";  
 RL Biochem. Biophys. Res. Commun. 237:297-302(1997).  
 RN [4]  
 RP ERRATUM.  
 RX MEDLINE=98069816; PubMed=9412455;  
 RA Maier R., Glatz A., Mosbacher J., Bilbe G.;  
 RL Biochem. Biophys. Res. Commun. 240:298-302(1997).  
 CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR UDP > UTP > ATP. THE ACTIVITY

OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.

-!- SUBCELLULAR LOCATION: Integral membrane protein.

-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X97058; CAA65770.1; --  
 DR EMBL; U52464; AAB03572.1; --  
 DR EMBL; AF007892; AAB07113.1; --  
 DR EMBL; AF007891; AAB07112.1; --  
 DR HSSP; P34996; 1DDD.  
 DR GeneW; HGNC:8543; P2RY6.  
 DR MIM; 602451; --  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODPSN.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEPTOR\_F1\_1; FALSE\_NEG.  
 DR PROSITE; PS0262; G-PROTEIN\_RECEPTOR\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 27 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 28 48 1 (POTENTIAL).  
 FT DOMAIN 49 62 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 63 83 2 (POTENTIAL).  
 FT DOMAIN 84 101 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 102 122 3 (POTENTIAL).  
 FT DOMAIN 123 144 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 145 165 4 (POTENTIAL).  
 FT DOMAIN 166 194 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 195 215 5 (POTENTIAL).  
 FT DOMAIN 216 236 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 237 257 6 (POTENTIAL).  
 FT DOMAIN 258 280 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 281 303 7 (POTENTIAL).  
 FT DOMAIN 304 328 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 5 5 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT DISULFID 99 177 BY SIMILARITY.  
 FT CONFLICT 2 3 MISSING (IN REF. 2).  
 FT SEQUENCE 328 AA; 36429 MW; AAD6C55A43818107 CRC64;

Query Match 20.0%; Score 387.5; DB 1; Length 328;  
 Best Local Similarity 32.7%; Pred. No. 1.5e-18;  
 Matches 108; Conservative 42; Mismatches 143; Indels 37; Gaps 9;

Qy 33 DMNTSQEGL---CFSEKVKQVYLSLAYSTFIETLGLPLNGTIVLHWSWGOTKRWSCATT 88  
 Db 4 DNGTQALGLPTTCVYRENFQKLLPPVYSAVLAAGLPLNCIVTQICTSRRLTRAV 63  
 Qy 89 YLVNLMVADLLVYL-LPELIITYSLDRWPFGEELCKLVHFLYINLYGSLILLTFCISVH 147  
 Db 64 YTLNLADLLYACSLPLLYNVAQGDHWPFGDFACRLVRFELFYANLHGSILFLTCISFQ 123  
 Qy 148 QFLGVCHPLCSLPYR-TRHAWLGTSTTVALVVLQLLPTLAFSHDYINGQMIWDMTSQ 206  
 Db 124 RYLGIChPLAPWKKRGGRAWLVCVAVMLAVTTQCLPTAFAATGIQRNRTVCYDLSP 183  
 Qy 207 ENFDRLFAYGIVLTLSGFL---SLIGHGVLFTDQEPDQARGEPHEDRHSQVHPD 262  
 Db 184 ALATHYMPYGMALTVIGFLFPALACLLAC-----RLCRDGPAPV 231  
 Qy 263 HPTGVWPLHPLFCALPYHSLHLLPHLL-SAFSGPLPDGSCGLDMEASGEQLPOP- 320  
 Db 232 ERGRGAARMAVVAFAISFLPFHTKTAYLAVSTGVPCTV--LEAFAAAYKGRPF 289  
 Qy 321 -----SPVLSF-----KGGKNRVRLLQKL 339  
 Db 290 ASANSVLPDILFYFTQKKFRRRPHELLQKL 319

RESULT 7

P2Y2\_HUMAN

ID P2Y2\_HUMAN STANDARD; PRT; 377 AA.

AC P41231; Q96EM8;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)  
 DE (Purinergic receptor).  
 GN P2RY2 OR P2RUI.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Airway epithelium;  
 RX MEDLINE=94211846; PubMed=8159738;  
 RA Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H., Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;  
 RT "Cloning and expression of a human P2U nucleotide receptor, a target for cystic fibrosis pharmacotherapy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:3275-3279(1994).  
 RN [2]  
 RP REVISTONS  
 RX MEDLINE=95108098; PubMed=7809171;  
 RA Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H., Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;  
 RT "Cloning and expression of a human P2U nucleotide receptor, a target for cystic fibrosis pharmacotherapy.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:13067-13067(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney, and Leukocyte;  
 RA Strausberg R.;  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBSJ databases.  
 CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that activate a phosphatidylinositol-calcium second messenger system. The affinity range is UTP = ATP > ADP-gamma-S >> 2-methylthio-ATP = ADP.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: SPLEEN, TESTIS, KIDNEY, LIVER, LUNG, HEART AND BRAIN.  
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; U07225; AAC04923.1; --  
 DR EMBL; BC012104; AAH12104.1; --  
 DR EMBL; BC028135; AAH28135.1; --  
 DR PIR; A54946; A54946.  
 DR HSSP; P34996; 1DDD.  
 DR GeneW; HGNC:8541; P2RY2.  
 DR MIM; 600041; --  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODPSN.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEPTOR\_F1\_1; 1.  
 DR PROSITE; PS0262; G-PROTEIN\_RECEPTOR\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.  
 FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 33 59 1 (POTENTIAL).  
 FT DOMAIN 60 70 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 71 93 2 (POTENTIAL).  
 FT DOMAIN 94 110 EXTRACELLULAR (POTENTIAL).





FT	TRANSMEM	73	93
FT	DOMAIN	96	112
FT	TRANSMEM	113	131
FT	DOMAIN	132	154
FT	TRANSMEM	155	174
FT	DOMAIN	175	196
FT	TRANSMEM	197	222
FT	DOMAIN	223	246
FT	TRANSMEM	247	269
FT	DOMAIN	270	287
FT	TRANSMEM	288	309
FT	DOMAIN	310	365
FT	DISULFID	108	185
FT	VARIANT	168	168
FT	VARIANT	178	178
FT	VARIANT	191	191
FT	CONFLICT	86	86
FT	CONFLICT	234	234
SQ	SEQUENCE	365 AA;	40963 MW; 23E0AFED3B7DDEED CRC64;

Query Match 19.0%; Score 367; DB 1: Length 365;  
Best Local Similarity 41.1%; Pred.No. 3.4e-17;  
Matches 79; Conservative 30; Mismatches 75; Indels 8; Gaps

QY	43	CQFSERYKQVLSLAYSIIFILGLPLNGTVLWHSWGQTKRWSCATTYLVNLAMVADLLYYVL	102
DB	27	CWEDEFKFILLPVSAVVVFVLGSLNAPTFLWFIERLRPDATATYMFHLASDTLYVL	36
QY	103	-LPFLITYSLDRWPFGELLCKLVIHFIVINLYGSILLTLTCISVHQFLGVCHPCLSLPY	161
DB	87	SLRPTLIYYAAHNHPFEGTEICKEFRVEFYWNLYCSVLFITCISVHYRLGICHPLRALRW	146
QY	162	RTRRHAWLGTSSTWALVWLQLLTPLAFSHPTYINGOMIWDYDMT SQENDFRLFAYGI VLT	221
DB	147	GRPLRAGLICLAWLVLVAGCLVPNLFEVTISKNTTVLCHDTRPEEFDHVYHVS-----	201
QY	222	SGFSLLLGHGV	233
DB	202	SAVMGLL--PGV	211

RESULT 11  
P2Y2\_RAT STANDARD; PRT; 374 AA.

ID	P2Y2_RAT	STANDARD;	PRT;	374 AA.
AC	P41232;			
DC	01-FEB-1995 (Rel. 31, Created)			:
DT	01-OCT-1996 (Rel. 34, Last sequence update)			:
DT	15-JUN-2002 (Rel. 41, Last annotation update)			:
DE	P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)			:
DE	(Purinergic receptor).			:
GN	P2RY2 OR P2RU1.			:
OS	Rattus norvegicus (Rat).			:
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			:
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			:
OX	NCB Taxid=10116;			:
ON	[1]			:
RP	SEQUENCE FROM N.A.			:
RX	MEDLINE=95110548; PubMed=7811468;			:
RP	Rice W.R., Burton F.M., Fiedelvey D.T.;			:
RA	"Cloning and expression of the alveolar type II cell p2u-purinergic			:
RT	receptor.";			:
RL	Am. J. Respir. Cell Mol. Biol. 12:27-32(1995).			:
RL	[2]			:
RC	SEQUENCE FROM N.A.			:
RC	STRAIN=Wistar; TISSUE=Pituitary;			:
RA	Chen Z.P., Krull N., Xu S., Levy A., Lightman S.L.;			:
RL	Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.			:
RL	[3]			:
RC	SEQUENCE FROM N.A.			:
RC	STRAIN=Wistar;			:







Matches	77:	Conservative	36:	Mismatches	82:	Indels	15:	Gaps			
QY	53	YLSLAYSIIFILPLPLNGTVLWHSGQTKRWSCATTYLVNLMVADLLYVL-LPFLIITYS	111								
Db	53	YLPVAVILYFIIGFLGNSVAIMWFVHPKPSGISVYMENALADFLYVLTLPALIFYF	112								
QY	112	LDRNPFGEGLCKLVHPLFYINLYGSIILLTLCISVHQFLGVCHPLCSLPYRTRHAWLGT	171								
Db	113	NKTDWIFGDVCKLQRFIFHVNLYGSIILFTCTISAHRSYGVVYPLKSLGRKKKNAIYVS	172								
QY	172	STTWALVILQLLPLAFSHTDY-INGQMTWYDMTSGENDRFLFAYGIVTLGSGF----	226								
Db	173	VLVWLIVVAISILPYSYSGTGIRKNTVYCYDSTDEYLSRFIYSNCTTVAMFCIPLVL	232								
QY	227	LLGHFGVLETFDGOEPDQARPEHEDRQHP	256								
Db	233	ILGCYGLI-----VRALIYKDLNSP	253								
RESULT 15											
P2YR_HUMAN											
ID	P2YR_HUMAN	STANDARD;	PRT;	373	AA.						
AC	P47900;										
DT	01-FEB-1996 (Rel. 33, Created)										
DT	01-FEB-1996 (Rel. 33, Last sequence update)										
DT	15-JUN-2002 (Rel. 41, Last annotation update)										
DE	P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).										
GN	P2RY1.										
OS	Homo sapiens (Human).										
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;										
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.										
OX	NCBI_Taxid:9606;										
RP	[1]										
RP	SEQUENCE FROM N.A.										
RC	TISSUE=Placenta;										
RX	MEDLINE-96257237; PubMed-86666290;										
RA	Leon C., Vial C., Cazenave J.-P., Gachet C.;										
RT	"Cloning and sequencing of a human cDNA encoding endothelial P2Y1										
RT	purinoceptor.";										
RL	Gene 171:295-297(1996).										
RP	[2]										
RP	SEQUENCE FROM N.A.										
RX	MEDLINE-96158962; PubMed-8579591;										
RA	Ayyanathan K., Tania W., Harbansjit S., Raghbir A.S., Barnard E.A.,										
RA	Kunapuli S.P.;										
RT	"Cloning and chromosomal localization of the human P2Y1										
RT	purinoceptor.";										
RL	Biochem. Biophys. Res. Commun. 218:783-788(1996).										
RP	[3]										
RP	SEQUENCE FROM N.A.										
RX	MEDLINE-96205320; PubMed-8630005;										
RA	Janssens R., Communi D., Pirotton S., Samson M., Parmentier M.,										
RA	Boeynaems J.M.;										
RT	"Cloning and tissue distribution of the human P2Y1 receptor.";										
RL	Biochem. Biophys. Res. Commun. 221:588-593(1996).										
RP	[4]										

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- INDUCTION: REPRESSED BY THE P2Y1 RECEPTOR-SPECIFIC ANTAGONISTS  
CC A3P5P, A3P5P AND A2P5P. THESE INHIBIT CALCIUM ION MOBILIZATION  
CC AND SHAPE CHANGE IN PLATELETS.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----

DR EMBL; Z49205; CAA89066.1; -  
DR EMBL; U42030; AAA97873.1; -  
DR EMBL; U42029; AAA97872.1; -  
DR EMBL; S81950; AAB47091.1; -  
DR EMBL; AJ006945; CAA07339.1; -  
DR EMBL; AF018284; AAB94556.1; -  
DR HSSP; P34996; 1DDD.  
DR Genew; HGNC:8539; P2RY1.  
DR MIM: 601167; -

DR InterPro: IPR000276; GPCR\_Rhodopsn.

DR Pfam: PF00001; 7tm\_1; 1.

DR PRINTS: PR00237; GPCR\_Rhodopsn.

DR PROSITE: PS00237; G-PROTEIN\_RECEP.F1.1; 1.

DR PROSITE: PS0262; G-PROTEIN\_RECEP.F1.2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Platelet;

KW Blood coagulation.

FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 53 74 1 (POTENTIAL).

FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 88 109 2 (POTENTIAL).

FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 127 147 3 (POTENTIAL).

FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 167 188 4 (POTENTIAL).

FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 219 238 5 (POTENTIAL).

FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 266 285 6 (POTENTIAL).

FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 304 328 7 (POTENTIAL).

FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).

FT DISULFID 124 202 BY SIMILARITY.

FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 138 138 MISSING (IN REF. 1).

FT SEQUENCE 373 AA; 42071 MW; 4DC7C668B4145392 CRC64;

Query Match 16.9%; Score 328; DB 1; Length 373;

Best Local Similarity 29.9%; Pred No. 1.2e-14;

Matches 93; Conservative 51; Mismatches 123; Indels 44; Gaps 10;

QY 53 YLSLAYSIIFILGLPLNGTVLHWSGQTRKWSGATYLVNLMVADLYVL-LPFLIITYS 111

DB 53 YLPVAVILVFIIGFLGNSVAINMFVHMKPWSGISVYMFNLALADFLYVLTLPALIFYF 112

QY 112 LDDRWFPELLCKLVHFLYINLYGSIILLTCISVHQFVGVCHPLCSLPYRTRHAWLGT 171

DB 113 NKTDFWFGDAMCKLQRFIFHVNLYGSIILTCISAHRSQVYVYPLKSLGRLKKNAICIS 172

QY 172 STTWALVQLPLTFLAFSTHDYINGOMI-WYDMTQENFDRLFAVGIVLTLSGF----LS 226

DB 173 VLWLVVVAISILFYSGTVGRKNTICYDTSDEYLSRFIYSMCTTVMFCVPLVL 232

QY 227 LIGHGVLFDTGQEPDQARGEHPHQSPQVHPDPTGVNPLHPLFCALPYHSL-LLP 285

DB 233 ILGCYGLI-----VRALIVKLDLNSPLRK-----SIYLVIVLTVFAVSYIP 275

QY 286 HHLLSAFSGLPALD---GSQCGLQD-----MEASGECEQLPQSPVLSFKGGKN 331  
DB 276 FHVMTMNLRLARLDFTPAMCAFNDRVYATYQVTRGLASLNSC-----VDPILYFLAGDT 330  
QY 332 -RVRLLOKLRQ 341  
DB 331 FRRRLSRATRK 341

Search completed: May 22, 2003, 12:50:04  
Job time : 16 secs







DR	EMBL; AF254119;	AAF66063.1;	-			
DR	HSSP; P34996;	IDDD.				
DR	InterPro; IPR000276;	GPCR_Rhodpsn.				
DR	Pfam; PF00001;	7tm_1; 1.				
DR	PRINTS; PR00237;	GPCRRHODPSN.				
DR	PROSITE; PS00237;	G_PROTEIN_RECPE_FL1; 1.				
DR	PROSITE; PS50262;	G_PROTEIN_RECPE_FL2; 1.				
KW	G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.					
SQ	SEQUENCE 359 AA; 41062 MW; C07010EDB8110EB CRC64;					
Query Match 12.8%; Score 248.5; DB 6; Length 359;						
Best Local Similarity 31.5%; Pred. No. 1.7e-15;						
Matches	64; Conservative	41; Mismatches	87; Indels 11; Gaps			
4;						
QY	34	MNTSQEQL-----CQSEKYKOVYLSL--AYSIIIFILGLPLNGTVLHWSGQTNRWSC	85			
Dd	3	LNSSTEDGIKRIODCPKAGRHNYFIMIPLYSIIFVVGFGSLVVIVIFYMKLKTV	62			
QY	86	ATTYVLNMVADLLYL-LPELIITYSDDDRMPGEGELCKLVHLFYINVXGSILLTCI	144			
Dd	63	ASVELLNALADLCFLTLPLMAVTAYMEYRWPGENVLCIASGVSNFYASFVLTCL	122			
QY	145	SVHQFLGCVPCLSPYTRRHMALGSTTWALVVQLLPFLAHSHTDYINGOMIWDMT	204			
Dd	123	SIDRYLAIVHPMSRLRTMLAVKTCTIIWLAGLASLPTIIHRNVFFIENTNITVCAF	182			
QY	205	SQENFDRLFAYGIVLT--LSGF L 225				
Dd	183	HYEQNSTLPVGLGTKNILGFL 205				
RESULT 8						
ID	Q9GKP7	PRELIMINARY;	PRT; 315 AA.			
IC	O9GKP7					
DT	01-MAR-2001	(TrEMBLrel. 16, Created)				
DT	01-MAR-2001	(TrEMBLrel. 16, Last sequence update)				
DT	01-MAR-2002	(TrEMBLrel. 20, Last annotation update)				
DE	Somatostatin receptor 1 (Fragment).					
GN	SSTR1.					
OS	Sus scrofa (Pig).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OX	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.					
CC	NCBI_Taxid=9823;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=20556840; PubMed=11105212;					
RA	Marklund S., Tuggle C.K., Rothschild M.F.;					
RT	Mapping of the CYP1A, SSTR1 and TRF1 genes to pig chromosome 7q					
RF	refines the porcine-human comparative map.";					
RL	Anim. Genet. 31:318-321(2000).					
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY). -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS. EMBL; AF148990; AAC40780.1; -; InterPro; IPR000276; GPCR_Rhodpsn. Pfam; PF00001; 7tm_1; 1. PRINTS; PR00237; GPCRRHODPSN. PROSITE; PS00237; G_PROTEIN_RECPE_FL1; 1. PROSITE; PS50262; G_PROTEIN_RECPE_FL2; 1. KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane. FT NON_TER 1 FT NON_TER 315 315 SQ SEQUENCE 315 AA; 34355 MW; OB973038CA72A882 CRC64;					
Query Match 12.8%; Score 248; DB 6; Length 315;						
Best Local Similarity 32.9%; Pred. No. 1.7e-15;						
Matches	74; Conservative	41; Mismatches	100; Indels 10; Gaps			
6;						
QY	8	SRGRSGSR-RGALLRGASRDMEKVMDNTTSOEOGLQCFSKEYKQVYLSLAYSIIFLGL	66			
Dd	8	SCXSXRGPGA---GTADGMEEPGRNASQNCTLSE--GGQSAILISFIYSVVICVL 61				
QY	67	PLNGTVLHWSGQTKRWSCATTYLVNLVMAD-LLYVLLPFLLIITYSLDWRPFGELCKL	125			

SQL	SEQUENCE	359 AA;	41087 MW;	AF90BET721E911BA	CRC64;
	Query Match	12.5%;	Score 241.5;	DB 4;	Length 359;
	Best Local Similarity	30.5%;	Pred. No. 8.4e-15;		
	Matches	62;	Conservative 42;	Mismatches 88;	Indels 11; Gaps
QY	34 MNTSQEOGL-----CQFSKYYKQVLSL--AYSIIFILGPLNGTVLVHWSGQTWRSC 85	: :   :	: : : :	: :	: : :
Dd	3 LNSSTEDGIKRIOODCPKAGRHNYIFVMIDPTLYSIIIFVVGIFGNLSLVIVIFYMKLKV 62	:	:	:	:
QY	86 ATTIVNLWADLLYL-LPFLIITYSLDRWPFGELCKLVHLFFVINLYSGILLTLCI 144	: :   :	: :	: : : :	: :
Dd	63 ASVELLNLADCLFLLDPLNNAVYTAMETRWPFNGYNLKIASASVSFYASVELLTCL 122	:	:	:	:
QY	145 SVHOFGLGVCHPLGSPYTRRRHAWLGTTTWALVVLQLLPDLTAFTSDTYINGOMIWDMT 204	: : :	:   :	:   :	:   :
Dd	123 SIDRYLAIVHPMSRLRRTMLVAKVTCTIIWLLAGLASLPAITHRVVFIENTNITVCAP 192	: : :	:   :	:   :	:   :
QY	205 SQENFDRLFAYGIVLT--LSGFL 225	: : :	:   :	:   :	:   :
Dd	183 HYESQNSTLPIGLGLTKNILGFL 205	:	:	:	:
RESULT 12					
Q9DGII	ID	Q9DGII	PRELIMINARY;	PRT;	358 AA.
AC	Q9DGII;				
DT	01-MAR-2001 (TREMBLrel. 16, Created)				
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	Chemokine receptor CXCR4.				
GN	NCBI_TaxID=9031;				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;				
OC	GALLUS.				
OX	NCBI_TaxID=9031;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21089206; PubMed=11272281;				
RA	Liang T.S., Hartt J.K., Lu S., Martins-Green M., Gao J.-L.,				
RA	Murphy P.M.;				
RT	"Cloning, mRNA distribution, and functional expression of an avian				
RL	counterpart of the chemokine receptor/HIV coreceptor CXCR4.";				
RL	J. Leukoc. Biol. 69:297-305(2001).				
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).				
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.				
DR	EMBL; AF294794; AAC09054.1; -				
DR	InterPro; IPR000276; GPCR_Rhodpsn.				
DR	Pfam; PF00001; 7tm_1; 1.				
DR	PRINTS; PR01558; CHEMOKINER11.				
DR	PRINTS; PR00237; GPCRRHODPSN.				
DR	PROSITE; PS00237; G_PROTEIN_RECEP_Fl_1; 1.				
DR	PROSITE; PSS0262; G_PROTEIN_RECEP_Fl_2; 1.				
KW	G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.				
SQL	SEQUENCE	358 AA;	40094 MW;	64E4977E29C86C9E	CRC64;
	Query Match	12.1%;	Score 235;	DB 13;	Length 358;
	Best Local Similarity	27.3%;	Pred. No. 3.5e-14;		
	Matches	74;	Conservative 56;	Mismatches 113;	Indels 28; Gaps
QY	20 LLLEGASRDMEKVDMMNTSQEOGLCFQSEK---YKQVYLSLAYSIIFTGLPLNCTVLWHS 76	: :   :	: :	: : : :	: :
Dd	10 ILIEFANGSEEIGSADYGDEPCFOHEANDFRIPTYISIFTGLIGNGLVIIVM 69	:	:	:	:
QY	77 WGQTKRWSCATTYLVNLNMADLLYL-LPFLIITYSLDRWPFGELCKLVHLFFVINLY 135	: : :	:   :	:   :	:   :
Dd	70 GYQKKQRSMTDKYRLHLSVADLLFVITLPFWSDAAIS--WYEGNLCAKHVIYTVNLY 127	:	:	:	:
QY	136 GSILLLTCTISVHQFLGVCHPLCSLPYR---TRRHAWLGTTTWALVVQLLPDLTAFTSHD 192	: : :	:   :	:   :	:   :
Dd	128 SSVLILAFISLDRLYAIVHATNSQRPKLAEIKTVYG---VWLPVALLTPVIDIASTS 184	: : :	:   :	:   :	:   :

```
QY 193 YINGQMIWDMTQENFDRFAYG---IVLTLSGLSLGHFGVLTGQEPDQARGEPH 249
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21172992; PubMed=11273702;
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in
RT the discovery of new g-protein coupled receptors.";
RL J. Mol. Biol. 307:799-813(2001).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF348078; AAK29080.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 330 AA; 38283 MW; 6D3C70E340866BAB CRC64;

RESULT 13
Q9EQR9 PRELIMINARY; PRT; 359 AA.
AC Q9EQR9;
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Angiotensin II receptor type 1B.
GN AT1B.
OS Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Meriones.
OX NCBI_TaxID=10047;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=ADRENAL GLAND;
RA Hoe K.-L., Saavedra J.M.;
RT "Molecular cloning and pharmacological characterization of an atypical
RT gerbil angiotensin II type-1 receptor and amino acids which are
RT important in losartan binding.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF078794; AAG43172.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 359 AA; 40934 MW; 3FALCCE9B734A29 CRC64;

Query Match 12.0%; Score 233; DB 11; Length 359;
Best Local Similarity 28.1%; Pred. No. 5.4e-14;
Matches 65; Conservative 49; Mismatches 107; Indels 10; Gaps 4;

QY 35 NTSQEQGL-----QFSEKQKVYLSL--AYSIIFILGLPLNGTVLHWSGQTKRWSCA 86
DB 4 NSTEDGIRRIQDDCPKAGRHNYIFVMITPLYSIIIFVMGIFGNSLVVIYFYMKLKTVA 63
QY 87 TTYLVNLMVADLLYL-LPFLIITYSLDDRWPFGLLCKLVHFLFYINLYSGILLITCIS 145
DB 64 SVFLNLALADLCFLTLPLMAVYTAKEYRPFNGHLCCKIASAGVSFNLYASVELLTCL 123
QY 146 VHQFLGVCHPLCSLPYRTRRRHAWLTSTTVALVQLLPTLAFSHTDYINGOMIWDWTS 205
DB 124 IDRYLAIVHPRSLRRLMLVAKVTCIIVIIWLAGLASPAVTHRNVTNTNITVCAPIH 183
QY 206 QNFDRFLFAYGIVLTLSGLSLGHFGVLTGQEPDQARGEPHEDROHSP 256
DB 184 YESQNSTLPVGMGLT-KNLTGFFVFFLIILTSYTLIWKALKAYELQKNKP 233

RESULT 14
Q9BXA5 PRELIMINARY; PRT; 330 AA.
AC Q9BXA5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE G-protein coupled receptor 91.
GN GPR91.

QY 193 YINGQMIWDMTQENFDRFAYG---IVLTLSGLSLGHFGVLTGQEPDQARGEPH 249
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21172992; PubMed=11273702;
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in
RT the discovery of new g-protein coupled receptors.";
RL J. Mol. Biol. 307:799-813(2001).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF348078; AAK29080.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 330 AA; 38283 MW; 6D3C70E340866BAB CRC64;

Query Match 12.0%; Score 232.5; DB 4; Length 330;
Best Local Similarity 33.9%; Pred. No. 5.6e-14;
Matches 61; Conservative 34; Mismatches 80; Indels 5; Gaps 3;

QY 47 EYKQVYLSLAYSIIFILGLPLNGTVLHWSGQTKRWSGQTKRWSGQTKRWSGQTKRWSG 105
DB 22 EYK---YLSIFGIEFVGVGLNTIVVGYIFSLKNWSSNIYLFNLSVSDLAFLCTLPW 78
QY 106 LIITYSLDDRWPFGLLCKLVHFLFYINLYSGILLITCISVHQFLGVCHPLCSLPYRTR 165
DB 79 LIRSYA-NGNIYGDVLCISNRYVLHANLYTSTLFTFISIDRYLIKIYPFREHLQKKE 137
QY 166 HAWLGTSTTVALVQLLPTLAFSHTDYINGOMIWDWTSQENFDRFLFAYGIVLTLSGL 225
```



Db 138 FAILISLAIWVLTLELPILPLINPVITDNGTTCNDFASSGDPNXYNIYSMCLTLGLFL 197

Search completed: May 22, 2003, 12:50:46  
Job time : 36 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 26, 2003, 21:29:43 ; Search time 304 Seconds  
(without alignments)  
2666.840 Million cell updates/sec

Title: US-10-023-586B-2

Perfect score: 1936

Sequence: 1 MSLILLSRSGSRGRRGAL.....QNKLGHPAGKRCPLNRS 360

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-DB=/cgn2\_1/USPto\_spool/US10023586/Runat\_22052003\_114849\_28633/app\_query.fasta\_1.519  
-QB=N\_Geneseq\_101002 -QFMT=fastp -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10023586.ecgn\_1\_1\_263\_etunat\_22052003\_114849\_28633 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEOQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1936	100.0	3143	24	Human genomic clon
2	1898	98.0	1063	24	Human DNA for pote
3	1798	92.9	1002	22	Human G-protein-co
4	1785	92.2	1076	24	Human G-protein co
5	1541	79.6	850	24	Human P2Y-like G p
6	921	47.6	510	22	Human ngPCR12 codi
7	732	37.8	681	22	Human CDNA encodin
8	681	35.2	585	22	Human purinergic r
9	453	23.4	1163	22	Human P2Y recepto
10	412	22.8	1429	18	Human purinergic r
11	420	21.7	984	17	Mouse pancreas G-p
12	418	21.5	1842	16	Human P20 receptor
13	415	21.4	2138	24	Mus musculus GPCR
14	415	21.4	2138	24	Mus musculus GPCR
15	413	21.3	984	17	Human placenta G-p
16	402	20.8	984	18	P2U purinergic rec
17	402	20.8	984	18	Human placenta pur
18	370	19.1	551	22	Human G-protein-co
19	370	19.1	551	22	Human G-protein-co
20	328	16.9	6721	24	Purinergic recepto
21	327	16.9	6721	24	Purinergic recepto
22	294	15.2	450	21	Human secreted exp
23	292	15.1	234	22	Human CDNA encodin
24	281	14.5	1020	22	Human ngPCR54 codi
25	278	14.4	1011	24	Human G-protein-co
26	278	14.4	1011	24	Human CDNA encodin
27	278	14.4	1014	24	Human G-protein co
28	278	14.4	1014	24	Human P2Y-like rec
29	278	14.4	1014	24	Human AXOR89 (G-pr
30	278	14.4	1014	24	Human purinergic-r
31	278	14.4	1014	24	Human G-protein co
32	278	14.4	1288	24	Human P2Y1-li enco
33	278	14.4	1288	24	Human CDNA encodin
34	278	14.4	1729	23	Human prostate exp
35	278	14.4	1729	23	Human prostate exp
36	278	14.4	1729	23	Human prostate exp
37	278	14.4	1729	23	Human prostate exp
38	278	14.4	1729	23	Human P2Y1-li enco
39	278	14.4	5435	24	Human purinergic-r
40	278	14.4	9905	24	Human polynucleoti
41	278	14.4	1313	22	Human G-protein co
42	275	14.2	1301	17	Human G-protein co
43	275	14.2	1301	21	Human G-protein co
44	275	14.2	1301	21	Human G-protein co
45	269	13.9	2162	19	Human mu-opioid re

ALIGNMENTS

RESULT 1

AAS17746

ID AAS17746 standard; DNA; 3143 BP.

XX AAS17746;

AC AAS17746;

DT 26-FEB-2002 (first entry)

XX Human genomic clone for P2Y-like G protein-coupled receptor..

DE Human; ds; P2Y-like G protein-coupled receptor; GPCR; COPD;

XX Chronic obstructive pulmonary disease; nervous system disease;

KW Parkinson's disease; multiple sclerosis; dementia; stroke;

KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;

KW bacterial infection; fungal infection; protozoan infection;

KW viral infection; pain; cancer; anorexia; bulimia; asthma;

KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;

KW angina pectoris; myocardial infarction; ulcer; inflammation; allergy;



psychotic disorder; neurological disorder; anxiety; schizophrenia;  
manic depression; delirium; severe mental retardation; dyskinesia.

Homo sapiens.

Key Location/Qualifiers  
CDS 520..2498

/\*tag= a  
/product= "P2Y-like GPCR"

W0200185764-A2.

15-NOV-2001.

09-MAY-2001: 2001WO-EP05244.

11-MAY-2000: 2000US-203582P.

21-FEB-2001: 2001US-269857P.

(FARB ) BAYER AG.

Ramakrishnan S;

WPI: 2002-075242/10.

P-PSDB; AAU11251.

New polynucleotides for producing P2Y-like G protein-coupled receptors (GPCR) that are used for screening inhibitors or regulators of human P2Y-like GPCR, especially useful for treating pain, cancer or neurological disorders

Disclosure, Fig 1: 114pp; English.

The invention relates to an isolated polynucleotide encoding a P2Y-like G protein-coupled receptor (GPCR) polypeptide, its fragment, derivative or allele, a host cell containing an expression vector comprising the polynucleotide and screening for agents that regulate the GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR polypeptide, which may be employed for screening agents that inhibit or regulate human P2Y-like GPCR. The reagent or inhibitor of the human P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR disorders, particularly COPD (chronic obstructive pulmonary disease), peripheral or central nervous system disease (e.g. Parkinson's disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and many other diseases and disorders listed in the specification), benign prostatic hyperplasia or urinary incontinence. A pharmaceutical composition containing the modulators and/or regulators of P2Y-like GPCR is useful for modulating the activity of a P2Y-like GPCR.

In particular, these are useful for treating, preventing or ameliorating infections (e.g. bacterial, fungal, protozoan or viral infections), pain, cancer, anorexia, bulimia, asthma, acute heart failure, hypotension, hypertension, osteoporosis, diabetes, angina pectoris, myocardial infarction, ulcers, inflammation, allergies, psychotic or neurological disorders (e.g. anxiety, schizophrenia, manic depression, delirium, severe mental retardation or dyskinesias). The present sequence is a genomic clone encoding the P2Y-like GPCR of the invention.

Sequence 3143 BP; 818 A; 771 C; 726 G; 828 T; 0 other;

#### Alignment Scores:

Pred. No.:	1-29e-174	Length:	3143
Score:	1936.00	Matches:	360
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	24	Gaps:	0

US-10-023-586B-2 (1-360) x AAS17746 (1-3143)

QY 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
DB 436 ATGCTGTCCATTTCCTCCAGGGAGACAGAGGGAGCGCTGAGCTCTG 495

QY 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluIn 40  
DB 496 CTCTGGAGGAGGCTCCCGGACATGGAGAGGTGGACATGAATACATCAGAGCAAA 555  
QY 41 GlyLeuCysGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60  
DB 556 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGAAAGTCTACCTCTCCCTGGCCCTACAGTATC 615  
QY 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrPheHisSerTrpGlyGlnThr 80  
DB 616 ATCTTTATCTAGGCTGCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGGCCAAACCC 675  
QY 81 LysArgTyrSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
DB 676 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCCACCTGCTTTAT 735  
QY 101 ValLeuLeuProPheLeuIleThrTyrSerLeuAspArgTrpProPheGlyGlu 120  
DB 736 GTGCTATTGCCCTTCTCATCATCACTACTACCTAGATGACAGGTGGCCCTTCGGGGAG 795  
QY 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
DB 796 CTGCTCTGCAAGCTGGTGACCTTCTCTGTATATCACTTTTACGGCAGCATCTCTGCTG 855  
QY 141 LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
DB 856 CTGACCTGATCTCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTGTGCTGCC 915  
QY 161 TyrArgThrArgHisAlaThrLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180  
DB 916 TACCGGACCGCAGGATGCTGCTGGTGGCACCAGCACCACTGGGCCCTTGGTGGCTC 975  
QY 181 GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200  
DB 976 CAGCTGCTGCCACACTGGCTTCTCCACACGACATACATCAATGGCCAGAGATGATCTGG 1035  
QY 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220  
DB 1036 TATGACATGACACCAAGAGAAATTTGATCGCTTTTGCCTACGGCATAAGTTCTGACA 1095  
QY 221 LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
DB 1096 TTGTCTGGCTTCTTCTCCCTCTGGTCAATTTTGGTGTCTATTTCACTGATGGTCAGGAG 1155  
QY 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260  
DB 1156 CCTGATCAAGCCAGAGAGAACCTCATGAGACAGGCAACACACAGCCCGAGCCAGTCCAT 1215  
QY 261 ProAspHisProThrGlyValTyrProLeuHisProLeuPheCysAlaLeuProTyrHis 280  
DB 1216 CCGGACCATCTACTGTGTGGCTTTCACCTCTGTTTGTGGCTTCCATATATCAC 1275  
QY 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300  
DB 1276 TCGTCTCTTCTACCTACCATCTCTCTTCTCTCTTCTCAGAGCTGCCAGCTCTTGATGGC 1335  
QY 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320  
DB 1336 AGCCAGTGTGGCTTACAGATATGGAGCCCTCTGTGAGTGTGAGCAGCTGCCCAACCC 1395  
QY 321 SerProValLeuSerPheLysGlyLysValAsnArgValArgLeuLeuGlnLysLeuArg 340  
DB 1396 AGTCTGTACTTTCTTCAAGGGGGGCAAAAATAGAGTCAGGCTCCTCCAGAAATGAGG 1455  
QY 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
DB 1456 CAGAACAAAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTGAACATCT 1515

#### RESULT 2

AAS98135  
ID AAS98135 standard; DNA; 1063 BP.  
XX  
AC AAS98135;



616 ATCTTTATCTAGGGTGGCCACTAAATAGGCACTGCTCTGTGGCACTCTCTGGGGCCAAACC 675  
 QY LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuLeuTyr 100  
 Db AAGCGCTGGAGCTGGCCACCACCTATCTGGTGAACCTGATGGTGGCGGACCTGCTTTAT 735  
 QY ValLeuLeuProPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120  
 Db GTGCTATTGGCTTCTCCATCATCATCACTACTACTACTAGTAGACAGGTGGCCCTTCGGGGAG 795  
 QY LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140  
 Db CTGCTGTGCAAGCTGGTGCATCTCTCTCTATATACACCTTACGGCAGCATCTGCTG 855  
 QY LeuThrCysIleSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160  
 Db CTGACCTGCATCTCTGTGCACAGTCTCTAGTGTGTGGCCACCATCTGTTCGTGCTGCC 915  
 QY TyrArgThrArgHisAlaTrpLeuGlyThrSerThrThrTrpAlaLeuValValLeu 180  
 Db TACGGAGCCCGACAGCATGCTGGTGGCCACCATCTGTGGCCCTGGTGGTCTCTC 975  
 QY GlnLeuLeuProThrLeuAlaPheSerHisThrAspTyrIleAsnGlyGlnMetIleTrp 200  
 Db CAGCTGTGCCCCACACTGCTCTCTCCACAGGACTATCATATGCGCAGATGATCTGG 1035  
 QY TyrAspMetThrSerGlnGlnAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220  
 Db TATGACATGACCCAGCAGAGAAATTTGATCGCTTTTGGCTACGGCATAGTCTTGACA 1095  
 QY LeuSerGlyPheLeuSerLeuLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240  
 Db TTGCTGTGGCTTTCTTCTCCCTCTGGTCACTTTGGTGTGCTATTCACTGATGTCAGGAG 1155  
 QY ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260  
 Db CCGGACCATCTACTGGTGTGGCTCTCTTCCACCTCTGTTTGTGCTTCCATATCAC 1275  
 QY SerLeuLeuLeuProHisLeuLeuSerAlaPheSerGlnLeuProAlaLeuAspGly 300  
 Db TCCTCTCTTCTACTCACCATCTGCTTCTGCTTCTCAGACTCCAGCTCTTGATGGC 1335  
 QY SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320  
 Db AGCCAGTGTGGCTACAGATATGAGGCTCTGCTGAGTGTGAGCAGCGCTCAACCC 1395  
 QY SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuLeuGlnLysLeuArg 340  
 Db AGTCTCTACTTTCTTCAAGGGGGGCAAAATAGACTAGCTAGCTCTTCCAGAAATGAGG 1455  
 QY GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360  
 Db CAGAACAAAGTTGGTGTGATCCTCAGCTCCAGCTGGGAGGAAGATGCCCGGTTGACAGATCT 1515

RESULT 4  
 AC026960/c  
 LOCUS AC026960 164502 bp DNA linear HTG 13-APR-2000  
 DEFINITION Homo sapiens chromosome 3 clone RP11-170K4 map 3, WORKING DRAFT  
 SEQUENCE, 15 unordered pieces.  
 AC026960  
 VERSION AC026960.2 GI:7547202  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 164502)  
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 TITLE Homo sapiens chromosome 3, clone RP11-170K4

# JOURNAL REFERENCE AUTHORS

Unpublished  
 2 (bases 1 to 164502)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
 Boguslavsky,L., Boukhgaiter,B., Brown,A., Burkett,G.,  
 Campoliano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
 Collymore,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,  
 Dodge,S., Domno,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
 Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,  
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
 Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehocsky,J.,  
 Levine,R., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,  
 Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,  
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Rogov,P., Spencer,B.,  
 Pisanic,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
 Roy,A., Santos,R., Schauer,S., Severy,P., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Stange-Thomann,N., Tirrell,A., Travers,M., Trigilio,J.,  
 Tesfaye,S., Theodore,J., Tirrell,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
 Direct Submission  
 Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 13, 2000 this sequence version replaced gi:7328839.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L8491

Center clone name: 170\_K\_4

Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156927 bases at least Q40

Consensus quality: 160376 bases at least Q30

Consensus quality: 161800 bases at least Q20

Insert size: 170000; agarose-fp

Insert size: 163102; sum-of-contigs

Quality coverage: 4.4 in Q20 bases; agarose-fp

Quality coverage: 4.5 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently

consists of 15 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1

2917: contig of 2917 bp in length

2918 3017: gap of 100 bp

3018 6310: contig of 3293 bp in length

6311 6410: gap of 100 bp

6411 9447: contig of 3037 bp in length

9448 9547: gap of 100 bp

9548 14470: contig of 4923 bp in length

14471 14570: gap of 100 bp

14571 20528: contig of 5958 bp in length

20529 20628: gap of 100 bp

20629 24667: contig of 4039 bp in length

24668 24767: gap of 100 bp

24768 29367: contig of 4600 bp in length

29368 29467: gap of 100 bp

29468 36237: contig of 6770 bp in length









```

repeat_region      /rpt_family="AT_rich"
7555..7717
/rpt_family="MERSA"
complement(7788..8198)
/rpt_family="MLTII"
complement(8536..8591)
/rpt_family="L2"
9634..9936
/rpt_family="AluSx"
complement(10766..10814)
/rpt_family="MIR"
complement(10848..10964)
/rpt_family="MIR"
17399..17431
/rpt_family="TTAAA"n"
complement(17432..17546)
/rpt_family="MIR"
17637..18073
/rpt_family="Charlieb"
18278..18313
/rpt_family="(TG)n"
21428..21841
/rpt_family="MSTD"
complement(21847..21926)
/rpt_family="LTR33A"
22387..22631
/rpt_family="MLTIG1"
complement(22768..23212)
/rpt_family="Tigger3(Golem)"
complement(23222..23279)
/rpt_family="Tigger3(Golem)"
23311..23437
/rpt_family="MLTIE2"
24105..24225
/rpt_family="MIR"
24226..24517
/rpt_family="MER33"
24518..24529
/rpt_family="MIR"
complement(24639..24724)
/rpt_family="L2"
25420..25994
/rpt_family="L2"
26024..26059
/rpt_family="(TG)n"
26585..27405
/rpt_family="LMB8"
complement(27789..27855)
/rpt_family="L2"
28628..28778
/rpt_family="MER77"

Alignment Scores:
Pred. No.:      2.75e-163      Length:      167084
Score:          1936.00      Matches:      360
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              9              Gaps:        0

US-10-023-586B-2 (1-360) x AC092999 (1-167084)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgArgGlyAlaLeu 20
|||||
Db 160921 ATGCTGTCATTTTCTCTCCAGGGAAGCAGAGAGCGGCGGCTGGAGCTCTG 160980

Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40
|||||
Db 160981 CTCTGGAGGAGCGCTCCCGGACATGGAGAGGTGACATGAATCATCATCAGGAACAA 161040

Qy 41 GlyLeuGlnPheSerGluLysTyrLysGlnValTyrLeuSerLeuAlaTyrSerIle 60
|||||
Db 161041 GGTCTGTCCAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 161100

```

```

Qy 61 IlePheIleLeuGlyLeuProLeuAsnGlyThrValLeuTyrHisSerTrpGlyGlnThr 80
|||||
Db 161101 ATCTTTATCTAGGCTGCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGCCAAACC 161160

Qy 81 LysArgTrpSerCysAlaThrThrTyrLeuValAsnLeuMetValAlaAspLeuTyr 100
|||||
Db 161161 AAGCGTGGAGCTGTGCACACCACTATCTGTGAACCTGTGTGGCGGACCTGCTTTAT 161220

Qy 101 ValLeuLeuProPheLeuIleIleThrTyrSerLeuAspAspArgTyrPropheGlyGlu 120
|||||
Db 161221 GGTATTTGCCCTTCTCATCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 161280

Qy 121 LeuLeuCysLysLeuValHisPheLeuPheTyrIleAsnLeuTyrGlySerIleLeuLeu 140
|||||
Db 161281 CTGCTCTCAAGCTGGTGCACTTCTGTCTATATCAACCTTTACGGCAGCATCTGCTG 161340

Qy 141 LeuThrCysLysLeuSerValHisGlnPheLeuGlyValCysHisProLeuCysSerLeuPro 160
|||||
Db 161341 CTGACCTGCATCTGTGCACCACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 161400

Qy 161 TyrArgThrArgArgHisAlaTyrLeuGlyThrSerThrThrTyrAlaLeuValValLeu 180
|||||
Db 161401 TACCGGACCCGAGGCACTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 161460

Qy 181 GlnLeuLeuProThrLeuAlaGlySerHisThrAspTyrIleAsnGlyGlnMetIleTyr 200
|||||
Db 161461 CAGCTGTCTGCCACACTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 161520

Qy 201 TyrAspMetThrSerGlnGluAsnPheAspArgLeuPheAlaTyrGlyIleValLeuThr 220
|||||
Db 161521 TATGACATGACCAAGCAAGAAATTTTGTGCGGCTTTTGTGCGGCTGTGTGTGT 161580

Qy 221 LeuSerGlyPheLeuSerLeuGlyHisPheGlyValLeuPheThrAspGlyGlnGlu 240
|||||
Db 161581 TTGCTGTGCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 161640

Qy 241 ProAspGlnAlaArgGlyGluProHisGluAspArgGlnHisSerProSerGlnValHis 260
|||||
Db 161641 CCTGATCAAGCCAGAGAGAACTCTCATGAGGACAGCAACACAGCCGAGGAGTCCAT 161700

Qy 261 ProAspHisProThrGlyValTyrProLeuHisProLeuPheValAlaLeuProTyrHis 280
|||||
Db 161701 CCGGACCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 161760

Qy 281 SerLeuLeuLeuProHisHisLeuLeuSerAlaPheSerGlyLeuProAlaLeuAspGly 300
|||||
Db 161761 TCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 161820

Qy 301 SerGlnCysGlyLeuGlnAspMetGluAlaSerGlyGluCysGluGlnLeuProGlnPro 320
|||||
Db 161821 AGCCAGTGTGGCTTACAAGATATGGAGGCTCTGTGTGTGTGTGTGTGTGTGTGTGT 161880

Qy 321 SerProValLeuSerPheLysGlyGlyLysAsnArgValArgLeuGlnLysLeuArg 340
|||||
Db 161881 AGTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAAGGCTCTCTCCAGAAACCTGAGG 161940

Qy 341 GlnAsnLysLeuGlyGluHisProAlaGlyArgLysArgCysProGlyLeuAsnArgSer 360
|||||
Db 161941 CAGAACAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGGTGTGACAGATCT 162080

RESULT 6
AC021773      163958 bp      DNA      linear      HTG 10-SEP-2000
LOCUS      Homo sapiens clone RP11-12N17, WORKING DRAFT SEQUENCE, 15 unordered
DEFINITION      pieces.
ACCESSION      AC021773
VERSION      AC021773.6 GI:10047750
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

```



# REFERENCE AUTHORS JOURNAL REFERENCE AUTHORS

1 (bases 1 to 163958)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome, clone RP11-12N17  
Unpublished  
2 (bases 1 to 163958)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Becker, R., Bede, F.,  
Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G., Castle, A.,  
Choepe, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P.,  
Dearellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,  
Ferraira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,  
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,  
Macdonald, P., Marquis, J., McEwan, P., McGurk, A., McKernan, K.,  
McPheters, R., Melidrim, J., Meneus, L., Morrow, J., Naylor, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K.,  
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 10, 2000 this sequence version replaced gi:7248967.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L3302  
Center clone name: 12\_N17  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 152766 bases at least Q40  
Consensus quality: 158075 bases at least Q30  
Consensus quality: 160495 bases at least Q20  
Insert size: 174000; agarose-ff  
Insert size: 162558; sum-of-contigs  
Quality coverage: 4.5 in Q20 bases; agarose-ff  
Quality coverage: 4.9 in Q20 bases; sum-of-contigs

## TITLE JOURNAL COMMENT

NOTE: This is a 'working draft' sequence. It currently  
consists of 15 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 7254: contig of 7254 bp in length  
7255 7354: gap of 100 bp  
7355 8612: contig of 1258 bp in length  
8613 8712: gap of 100 bp  
8713 11463: contig of 2751 bp in length  
11464 11563: gap of 100 bp  
11564 14507: contig of 2944 bp in length  
14508 14607: gap of 100 bp  
14608 22452: contig of 7845 bp in length  
22453 22552: gap of 100 bp  
22553 35086: contig of 12534 bp in length  
35087 35186: gap of 100 bp  
35187 44278: contig of 9092 bp in length  
44279 44378: gap of 100 bp  
44379 56786: contig of 12408 bp in length  
56787 56886: gap of 100 bp

\* 56887 65270: contig of 8384 bp in length  
\* 65271 65370: gap of 100 bp  
\* 65371 76394: contig of 11024 bp in length  
\* 76395 76494: gap of 100 bp  
\* 76495 100320: contig of 23826 bp in length  
\* 100321 100420: gap of 100 bp  
\* 100421 113127: contig of 12707 bp in length  
\* 113128 113227: gap of 100 bp  
\* 113228 126756: contig of 13529 bp in length  
\* 126757 126856: gap of 100 bp  
\* 126857 151061: contig of 24205 bp in length  
\* 151062 151161: gap of 100 bp  
\* 151162 163958: contig of 12797 bp in length.

### FEATURES

source  
1. 163958  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-12N17"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. 7254  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
7355. 8612  
/note="assembly\_fragment"  
8713. 11463  
/note="assembly\_fragment"  
11564. 14507  
/note="assembly\_fragment"  
14608. 22452  
/note="assembly\_fragment"  
22553. 35086  
/note="assembly\_fragment"  
35187. 44278  
/note="assembly\_fragment"  
44379. 56786  
/note="assembly\_fragment"  
56887. 65270  
/note="assembly\_fragment"  
65371. 76394  
/note="assembly\_fragment"  
76495. 100320  
/note="assembly\_fragment"  
100421. 113127  
/note="assembly\_fragment"  
113228. 126756  
/note="assembly\_fragment"  
126857. 151061  
/note="assembly\_fragment"  
151162. 163958  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right

misc\_feature  
49230 a 32369 c 32422 g 48536 t 1401 others  
ORIGIN

Alignment Scores:  
Pred. No.: 4,05e-162 Length: 163958  
Score: 1923.00 Matches: 360  
Percent Similarity: 99.72% 100%  
Best Local Similarity: 99.72% Mismatches: 0  
Query Match: 99.33% Indels: 1  
DB: 2 Gaps: 0

US-10-023-586B-2 (1-360) x AC021773 (1-163958)

Qy 1 MetLeuSerIleLeuLeuProSerArgGlySerArgSerGlySerArgGlyAlaLeu 20  
Db 98103 ATGCTGTCATTTGCTTCCTTCCAGGGGAAACGACGAGCGCGTGGAGCTCTG 98162  
Qy 21 LeuLeuGluGlyAlaSerArgAspMetGluLysValAspMetAsnThrSerGlnGluGln 40  
Db 98163 CTCCTGGAGGAGCGCTCCCGGACATGGAGAGGTGGACATGATACATCAGAGGACAA 98222



<b>AF411108</b>					
<b>VERSION AF411108.1 GI:16566321</b>					
<b>KEYWORDS</b>					
<b>SOURCE</b>					
<b>ORGANISM</b>					
Homo sapiens.					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
<b>REFERENCE</b>					
<b>AUTHORS</b>					
1 (bases 1 to 1019)					
Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,					
Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.					
<b>TITLE</b>					
Discovery and mapping of ten novel G protein-coupled receptor genes					
<b>JOURNAL</b>					
Gene 275 (1), 83-91 (2001)					
<b>MEDLINE</b>					
21458557					
<b>PUBMED</b>					
11574155					
<b>REFERENCE</b>					
<b>AUTHORS</b>					
2 (bases 1 to 1019)					
Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,					
Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.					
<b>TITLE</b>					
Direct Submission					
<b>JOURNAL</b>					
Submitted (17-AUG-2001) Department of Pharmacology, University of					
Toronto, 8 Raddle Creek Rd. Rm. 4353, Toronto, Ont M5S 1A8, Canada					
<b>FEATURES</b>					
source					
1..1019					
Location/Qualifiers					
/organism="Homo sapiens"					
/db_xref="taxon:9606"					
1..1019					
gene="GPR79"					
pseudo					
BASE COUNT 200 A 308 C 252 G 259 T					
ORIGIN					
Alignment Scores:					
Pred. No.: 2,06e-154 Length: 1019					
Score: 1808.00 Matches: 339					
Percent Similarity: 99.7% Conservative: 0					
Best Local Similarity: 99.7% Mismatches: 0					
Query Match: 93.29% Indels: 0					
DB: 9 Gaps: 0					
US-10-023-586B-2 (1-360) x AF411108 (1-1019)					
Qy	1	MetLeuSerIleLeuLeuProSerATGGlySerArgSerGlySerArgAGlyValaleu	20		
Db	1	ATGCTGCCATTTCCTTCCAGCGGAGCAGACGGGAGCCGTCTGGAGCTCTG	60		
Qy	21	LeuLeuGlugLyAlaSerArgAspMetGLylsValAspMetAsnthrSerGlnGluGln	40		
Db	61	CTCCTGGAGGAGGCTCCCGGCACATGGAGAGGTGCACATGAATACATCACAGAACA	120		
Qy	41	GlyLeuCysGlnPheSerGluLysTyrlsGlnValtyrLeuSerLeuAlatySerile	60		
Db	121	GGTCTCGCCAGTTCTCAGAGAAGTAGCAAGCAAGTATACCTCTCCCTGGGCTACAGTATC	180		
Qy	61	IlePheIleLeuGlyLeuProLeuAsnGlyThrVallLeuTrpHisSerTrpGlyGlnThr	80		
Db	181	ATCTTTATCTAGGCTGCCACATAATGGCACACTGCTTGAGGCACTCTCTGGGCAACC	240		
Qy	81	LysArgTrpSerCysAlaThrThrtYrLeuValasnLeuMetValalaAspLeuLeuty	100		
Db	241	AAGCGCTGGAGCTGCCACCATCTCTGGTAACCTGATGTTGGCCGACCTGCTTTAT	300		
Qy	101	VallLeuLeuProPheLeuIlelThrtYrSerLeuAspArgthrProPheGlyglu	120		
Db	301	GTGCTATGTGCCCCCTCTCATCATCACCTACTAGTACAGGTGGCTTCTGGGGAG	360		
Qy	121	LeuLeuCysLysLeuValHisPheLeuPheTyrlleasnlLeutyrGlySerileLeuLeu	140		
Db	361	CTGCTCTGAAGCTGGGACACTTCCCTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTG	420		
Qy	141	LeuThrCysileSerValHisGlnPheLeuGlyValCysHisProLeuCysserLeupro	160		
Db	421	CTGACCTGCATCTCTGTGCACCAAGTCTCCTAGGTGTGTGCCCACTGTCTGCTGCC	480		
Qy	161	TyrArgThrArgArgHisAlatrpLeuGlyThrSerThrThrtrpAlaLeuValleu	180		





AX365511	AX365511	3143 bp	DNA	linear	PAT 21-FEB-2002
LOCUS	Sequence 1 from Patent WO0185764.				
DEFINITION	AX365511				
ACCESSION	AX365511				
VERSION	AX365511.1	GI:18873666			
KEYWORDS	human,				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1				
AUTHORS	Ramakrishnan,S.				
TITLE	Regulation of human p2y-like g protein-coupled receptor				
JOURNAL	Patent: WO 0185764-A, 15-NOV-2001;				
FEATURES	Bayer Aktiengesellschaft (DE)				
source	Location/Qualifiers				
	1..3143				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	520..>1374				
	/note="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAB24312.1"				
	/db_xref="GI:18873667"				
	/translation="MEKVDNMTSQEGLQCFSEKYQVYLSLAYSIIFILGLPLNGTVLWHSWGTRKMSCATYLVNMLVADLLYLLPLITYISDDRRPFCELCCKLVHFLFXYNLGSIILLTCLSHVQELGVCHPLCSLPYRERRHAWLGSTWTALVVLQQLPTLAFSHDYINGOMIWTMDMSQENFDRFLPAGIVLTLGSLFLSHGHFVLTQCEPDDQARGEPHEDRQHSPPVHPDHPHTGVPWPLPFLFCALPYHSLLLPHLLLSAFSGLPALDDGSCGGLODMEASGE"				
BASE COUNT	818 a	726 g	828 t		
ORIGIN					
Query Match	100.0%;	Score 1082;	DB 6;	Length 3143;	
Best Local Similarity	100.0%;	Pred. No. 1.7e-268;			
Matches 1082;	Conservative	0;	Mismatches	0;	Gaps
Qy	1	ATGCTGTCCATTTCCTTCCTCCAGGGGAAGCAGAGCGGAGCGCTGCTGGAGCTCTG	60		
Db	436	ATGCTGTCCATTTCCTTCCTTCCTCCAGGGGAAGCAGAGCGGAGCGCTGCTGGAGCTCTG	495		
Qy	61	CTCCTGGAGGAGCCTCCCGGACATGAGAAAGTGGACATGAATACATCACAGGAACAA	120		
Db	496	CTCCTGGAGGAGCCTCCCGGACATGAGAAAGTGGACATGAATACATCACAGGAACAA	555		
Qy	121	GGTCTCTGCCAGTTCACAGAAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC	180		
Db	556	GGTCTCTGCCAGTTCACAGAAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC	615		
Qy	181	ATCTTTATCTTAGGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC	240		
Db	616	ATCTTTATCTTAGGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC	675		
Qy	241	AAGCGCTGGAGCTGTGCCACCACTATCTGTGAACCTGTAGTGGTGGCGACCTGCTTTAT	300		
Db	676	AAGCGCTGGAGCTGTGCCACCACTATCTGTGAACCTGTAGTGGTGGCGACCTGCTTTAT	735		
Qy	301	GTGCTATTGCCCTTCCTCATCATCACCTACTACAGTAGATGACAGTGGCCCTTCGGGGAG	360		
Db	736	GTGCTATTGCCCTTCCTCATCATCACCTACTACAGTAGATGACAGTGGCCCTTCGGGGAG	795		
Qy	361	CTGCTCTCAAGCTGGTGCACTTCCTGTCTTATATCAACCTTTACGGCAGCATCTCGCTG	420		
Db	796	CTGCTCTCAAGCTGGTGCACTTCCTGTCTTATATCAACCTTTACGGCAGCATCTCGCTG	855		
Qy	421	CTGACCTGCATCTCTGTGCCACAGTTCCTAGTGTGTGCCACCACTGTGTTGCGTGCCTC	480		
Db	856	CTGACCTGCATCTCTGTGCCACAGTTCCTAGTGTGTGCCACCACTGTGTTGCGTGCCTC	915		
Qy	481	TACGGGACCGGAGCATGCTGGCTGGGACACAGCACCACTGGGCCCTTGGTGTCTCTC	540		
Db	916	TACGGGACCGGAGCATGCTGGCTGGGACACAGCACCACTGGGCCCTTGGTGTCTCTC	975		

25



```
QY 541 CAGCTGCTGCCACACTGCGCTTCTCCACACGAGTACATCAATGCCAGATGATCTGG 600
Db 976 CAGCTGCTGCCACACTGCGCTTCTCCACACGAGTACATCAATGCCAGATGATCTGG 1035
QY 601 TATGACATGACAGCCAGAGAAATTTGATCGGCTTTTCCCTACGCGCATAGTCTGACA 660
Db 1036 TATGACATGACAGCCAGAGAAATTTGATCGGCTTTTCCCTACGCGCATAGTCTGACA 1095
QY 661 TTGCTGCTGCTTTCTTCCCTCCCTGCTGCTATTTGGTGTGCTATTCTACTGATGTCAGGAG 720
Db 1096 TTGCTGCTGCTTTCTTCCCTCCCTGCTGCTATTTGGTGTGCTATTCTACTGATGTCAGGAG 1155
QY 721 CTTGATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCGAGCCAGGTCAT 780
Db 1156 CTTGATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCGAGGTCAT 1215
QY 781 CCGGACCATCTACTGCTGTGTCCTTCCACCTCTGTTTGTGCTCCCTCCATATCAC 840
Db 1216 CCGGACCATCTACTGCTGTGTCCTTCCACCTCTGTTTGTGCTCCCTCCATATCAC 1275
QY 841 TCGCTCCTTCTTACCTCACCATCTCTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGC 900
Db 1276 TCGCTCCTTCTTACCTCACCATCTCTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGC 1335
QY 901 AGCCAGTGTGGCTTACAGATATGAGGAGGCTCTGCTGAGTGTGAGAGCTGCTCAACCC 960
Db 1336 AGCCAGTGTGGCTTACAGATATGAGGAGGCTCTGCTGAGTGTGAGAGCTGCTCAACCC 1395
QY 961 AGTCTCTACTTCTTTCAGGGGGGCAAAATAGAGTCAAGCTCCCTCCAGAACTGAGG 1020
Db 1396 AGTCTCTACTTCTTTCAGGGGGGCAAAATAGAGTCAAGCTCCCTCCAGAACTGAGG 1455
QY 1021 CAGAACAGTGTGGTGTGAGTCCAGCTGGGAGGAGAGATGCCAGGCTTGAACAGATCT 1080
Db 1456 CAGAACAGTGTGGTGTGAGTCCAGCTGGGAGGAGAGATGCCAGGCTTGAACAGATCT 1515
QY 1081 GG 1082
Db 1516 GG 1517

RESULT 4
AC026960/c
LOCUS
DEFINITION Homo sapiens chromosome 3 clone RP11-170K4 map 3, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
AC026960
VERSION AC026960.2 GI:7547202
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
RENC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164502)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glnde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Larocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
```

Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (25-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 13, 2000 this sequence version replaced gi:7328839.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L8491

Center clone name: 170\_K\_4

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156927 bases at least Q40

Consensus quality: 160376 bases at least Q30

Consensus quality: 161800 bases at least Q20

Insert size: 170000; agarose-fp

Insert size: 163102; sum-of-contigs

Quality coverage: 4.4 in Q20 bases; agarose-fp

Quality coverage: 4.5 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 15 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 2917: contig of 2917 bp in length  
\* 2918 3017: gap of 100 bp  
\* 3018 6310: contig of 3293 bp in length  
\* 6311 6410: gap of 100 bp  
\* 6411 9447: contig of 3037 bp in length  
\* 9448 9547: gap of 100 bp  
\* 9548 14470: contig of 4923 bp in length  
\* 14471 14570: gap of 100 bp  
\* 14571 20528: contig of 5958 bp in length  
\* 20529 20628: gap of 100 bp  
\* 20629 24667: contig of 4039 bp in length  
\* 24668 24767: gap of 100 bp  
\* 24768 29367: contig of 4600 bp in length  
\* 29368 29467: gap of 100 bp  
\* 29468 36237: contig of 6770 bp in length  
\* 36238 36337: gap of 100 bp  
\* 36338 46377: contig of 10040 bp in length  
\* 46378 46477: gap of 100 bp  
\* 46478 59361: contig of 12884 bp in length  
\* 59362 59461: gap of 100 bp  
\* 59462 75221: contig of 15760 bp in length  
\* 75222 75321: gap of 100 bp  
\* 75322 93371: contig of 18050 bp in length  
\* 93372 93471: gap of 100 bp  
\* 93472 117463: contig of 23992 bp in length  
\* 117464 117563: gap of 100 bp  
\* 117564 137715: contig of 20152 bp in length  
\* 137716 137815: gap of 100 bp  
\* 137816 164502: contig of 26687 bp in length.

Location/Qualifiers

1. 164502

/organism="Homo sapiens"

FEATURES  
source

TITLE  
JOURNAL

COMMENT



```
/db_xref="taxon:9606"
/chromosome="3"
/map="3"
/clone="RP11-170K4"
/clone_lib="RPC1-11 Human Male BAC"
1..2917
/note="assembly_fragment"
3018..6310
/note="assembly_fragment"
6411..9447
/note="assembly_fragment"
9548..14470
/note="assembly_fragment"
14571..20528
/note="assembly_fragment"
20629..24667
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
24768..29367
/note="assembly_fragment"
29468..36237
/note="assembly_fragment"
36338..46377
/note="assembly_fragment"
46478..59361
/note="assembly_fragment"
59462..75221
/note="assembly_fragment"
75322..93371
/note="assembly_fragment"
93472..117463
/note="assembly_fragment"
117564..137715
/note="assembly_fragment"
clone_end:17
vector_side:left"
137816..164502
/note="assembly_fragment"
BASE COUNT 48539 a 32273 c 33197 g 49088 t 1405 others
ORIGIN
Query Match 100.0%; Score 1082; DB 2; Length 164502;
Best Local Similarity 100.0%; Pred. No. 2.4e-268;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTGTCATTTGCTTCTCCAGGGGAGCAGAGCGGAGCGCTGCTGGAGCTCTG 60
DB 123728 ATGCTGTCATTTGCTTCTCCAGGGGAGCAGAGCGGAGCGGAGCGCTGCTGGAGCTCTG 123669
QY 61 CTCCTGGAGGAGCTCCCGGAGCATGAGAGGTGGACATGAATACATCACAGGAACAA 120
DB 123668 CTCCTGGAGGAGCTCCCGGAGCATGAGAGGTGGACATGAATACATCACAGGAACAA 123609
QY 121 GGTCTGTCAGTTCTCAGAGAGTACAGCAAGTCTACTCTCCCTGGCCTACATATC 180
DB 123608 GGTCTGTCAGTTCTCAGAGAGTACAGCAAGTCTACTCTCCCTGGCCTACATATC 123549
QY 181 ATCTTTATCTAGGCTGCCACTAAATGGCAGTCTCTTGTGGCAGTCTCCGGGCCAAACC 240
DB 123548 ATCTTTATCTAGGCTGCCACTAAATGGCAGTCTCTTGTGGCAGTCTCCGGGCCAAACC 123489
QY 241 AAGCGCTGGAGCTGTGCCACCATCTCTGTGGAACCTGATGTGGCGGACCTGCTTTAT 300
DB 123488 AAGCGCTGGAGCTGTGCCACCATCTCTGTGGAACCTGATGTGGCGGACCTGCTTTAT 123429
QY 301 GTGCTATGCGCTTCTCTATCATCATCTACTACTAGTACAGAGTGGCCCTTCGGGGAG 360
DB 123428 GTGCTATGCGCTTCTCTATCATCATCTACTACTAGTACAGAGTGGCCCTTCGGGGAG 123369
QY 361 CTGCTCTGCAAGCTGTGACCTTCTGTTCTATATCAACCTTTACGGCAGCATCTGCTG 420
DB 123368 CTGCTCTGCAAGCTGTGACCTTCTGTTCTATATCAACCTTTACGGCAGCATCTGCTG 123309
```

```
QY 421 CTGACCTGCATCTCTGTGCACCAAGTTCTTAGTGTGTGCCACCCACTGTGTTCGTGCC 480
DB 123308 CTGACCTGCATCTCTGTGCACCAAGTTCTTAGTGTGTGCCACCCACTGTGTTCGTGCC 123249
QY 481 TACCGGACCCGAGGAGCATGCTGGTGGGACCAACAGCAGCAGCTGGTGGTCTCTC 540
DB 123248 TACCGGACCCGAGGAGCATGCTGGTGGGACCAACAGCAGCAGCTGGTGGTCTCTC 123189
QY 541 CAGCTGTGCCACCACTGGCTTCTCCACAGGAGTACATCAATGGCCAGATGATCTGG 600
DB 123188 CAGCTGTGCCACCACTGGCTTCTCCACAGGAGTACATCAATGGCCAGATGATCTGG 123129
QY 601 TATGACATGACCAAGAGAAATTTTATCGGCTTTTTCACCGGAGTACATGTTCTGACA 660
DB 123128 TATGACATGACCAAGAGAAATTTTATCGGCTTTTTCACCGGAGTACATGTTCTGACA 123069
QY 661 TTGCTGTGGCTTTCTTCCCTCCTTGGTGTCTATTTTGGTGTCTATTTTCTGATGCTCAGGAG 720
DB 123068 TTGCTGTGGCTTTCTTCCCTCCTTGGTGTCTATTTTGGTGTCTATTTTCTGATGCTCAGGAG 123009
QY 721 CCTGATCAAGCCAGAGAGAACTCATGAGGACAGCAACAGCCGAGCCAGGTCAT 780
DB 123008 CCTGATCAAGCCAGAGAGAACTCATGAGGACAGCAACAGCCGAGCCAGGTCAT 122949
QY 781 CCGGACCATCTTACTGTGTGGCTCTTACCCCTCTTGTGTGGCCCTTCCATATCAT 840
DB 122948 CCGGACCATCTTACTGTGTGGCTCTTACCCCTCTTGTGTGGCCCTTCCATATCAT 122889
QY 841 TCGCTCTTCTACCTCACCATCTGCTTCTGCTTTCTCAGGAGTCCAGCTCTTGATGGC 900
DB 122888 TCGCTCTTCTACCTCACCATCTGCTTCTGCTTTCTCAGGAGTCCAGCTCTTGATGGC 122829
QY 901 AGCCAGTGTGGCTTACAGATATGAGGAGCTCTGTGTAGTGTGAGCAGTGTCCAAACC 960
DB 122828 AGCCAGTGTGGCTTACAGATATGAGGAGCTCTGTGTAGTGTGAGCAGTGTCCAAACC 122769
QY 961 AGCTCTGTACTTTTCTTCAAGGGGGGCAAAATAGAGTCTCTTCCAGAACTGAGG 1020
DB 122768 AGCTCTGTACTTTTCTTCAAGGGGGGCAAAATAGAGTCTCTTCCAGAACTGAGG 122709
QY 1021 CAGACAAAGTTGGTGTAGCATCCAGCTGGGAGAGAGATGCCAGGTTGACAGATCT 1080
DB 122708 CAGACAAAGTTGGTGTAGCATCCAGCTGGGAGAGAGATGCCAGGTTGACAGATCT 122649
QY 1081 GG 1082
DB 122648 GG 122647

RESULT 5
AC092999 Homo sapiens 167084 bp DNA linear PRI 30-JUL-2002
LOCUS Homo sapiens 3 BAC Rpl1-170K4 (Roswell Park Cancer Institute Human
DEFINITION BAC Library) complete sequence.
ACCESSION AC092999
VERSION AC092999.3 GI:19033407
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 167084)
AUTHORS Muzny D.M., Adams C., Adio-Oduola B., Ali-Osman F.R., Allen C.,
Alsbrooks S.L., Amaralungu H.C., Are J.R., Avelle M., Banks T.,
Barbaria J., Benton J., Binage K., Blankenburg K., Bonnin D.,
Bouck J., Bowie S., Brieva M., Brown E., Brown M., Bryant N.P.,
Buhay C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,
Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,
Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,
Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,
Davila M.L., Davis C., Davy Carroll L., Dederich D.A.,
Delaney K.R., Deigado O., Denn A.L., Ding Y., Dinh H.H.,
```



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 20:51:54 ; Search time 3164 Seconds  
(without alignments)  
9952.350 Million cell updates/sec

Title: US-10-023-586B-1  
Perfect score: 1082  
Sequence: 1 atgtgtccattttgtctcc.....ccagggtgaacagatctgg 1082

oring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*

- 1: gb\_ba:\*
- 2: gb\_htg:\*
- 3: gb\_in:\*
- 4: gb\_om:\*
- 5: gb\_ov:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*
- 15: em\_ba:\*
- 16: em\_fun:\*
- 17: em\_hum:\*
- 18: em\_in:\*
- 19: em\_mu:\*
- 20: em\_om:\*
- 21: em\_or:\*
- 22: em\_ov:\*
- 23: em\_pat:\*
- 24: em\_ph:\*
- 25: em\_pl:\*
- 26: em\_ro:\*
- 27: em\_sts:\*
- 28: em\_un:\*
- 29: em\_vl:\*
- 30: em\_htg\_hum:\*
- 31: em\_htg\_inv:\*
- 32: em\_htg\_other:\*
- 33: em\_htg\_mus:\*
- 34: em\_htg\_pln:\*
- 35: em\_htg\_rtd:\*
- 36: em\_htg\_mam:\*
- 37: em\_htg\_vrt:\*
- 38: em\_sy:\*
- 39: em\_hgo\_hum:\*
- 40: em\_hgo\_mus:\*
- 41: em\_hgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1082	100.0	1082	6	AX458306	AX458306 Sequence
2	1082	100.0	1334	9	AB065670	AB065670 Homo sapi
3	1082	100.0	3143	6	AX365511	AX365511 Sequence
c 4	1082	100.0	164502	2	AC026960	AC026960 Homo sapi
5	1082	100.0	167084	9	AC092999	AC092999 Homo sapi
6	1071	99.0	163958	2	AC021773	AC021773 Homo sapi
7	1007	93.1	1019	9	AF411108	AF411108 Homo sapi
8	1002.8	92.7	1020	6	AX458308	AX458308 Sequence
9	1002.8	92.7	1076	6	AX375230	AX375230 Sequence
10	850	78.6	850	6	AX365514	AX365514 Sequence
11	649.4	60.0	209273	2	AC121579	AC121579 Mus muscu
c 12	644.4	59.6	91107	2	AC116265	AC116265 Rattus no
13	510	47.1	510	6	AX147768	AX147768 Sequence
c 14	481.6	44.5	192524	2	AC120742	AC120742 Rattus no
15	396.4	36.6	681	6	AX244718	AX244718 Sequence
16	370	34.2	585	6	AX244727	AX244727 Sequence
17	235.8	21.8	1163	5	GDP2Y3	X98283 G.domesticu
18	227.8	21.1	1074	5	AF069555	AF069555 Meleagris
19	227.8	21.1	2025	9	HSU07225	U07225 Homo sapien
20	226.2	20.9	2257	9	BC028135	BC028135 Homo sapi
21	224.6	20.8	2681	9	BC012104	BC012104 Homo sapi
22	224.6	20.8	123185	2	AP000587	AP000587 Homo sapi
c 23	224.6	20.8	196988	9	AP002761	AP002761 Homo sapi
24	222	20.5	551	6	AX230145	AX230145 Sequence
25	203.8	18.8	1842	6	I34573	I34573 Sequence 1
26	203.8	18.8	1842	6	I36524	I36524 Sequence 1
27	203.8	18.8	1842	6	I76280	I76280 Sequence 1
28	202.2	18.7	2138	6	AX454153	AX454153 Sequence
29	202.2	18.7	2562	10	MUSP20REC	L14751 Mouse ATP r
30	202.2	18.7	2562	10	BC006613	BC006613 Mus muscu
31	201.4	18.6	1977	5	AF031897	AF031897 Meleagris
32	200.6	18.5	175947	2	AC079641	AC079641 Mus muscu
33	199	18.4	1131	10	S83099	S83099 P2U recepto
34	187.8	17.4	1688	10	RNU56839	U56839 Rattus norv
35	187.8	17.4	1918	10	RNU09402	U09402 Rattus norv
36	187.8	17.4	2011	10	RATP2U	L46865 Rattus norv
37	178.6	16.5	1232	9	HSGPCRE	X96597 H.sapiens g
38	178.6	16.5	1429	9	A62985	A62985 Sequence 1
39	178.6	16.5	1429	9	HSP2YA	X91852 H.sapiens P
c 40	178.6	16.5	178868	9	AL357752	AL357752 Human DNA
41	177	16.4	178861	2	AC068231	AC068231 Homo sapi
42	173.8	16.1	1651	9	HSU40223	U40223 Human uridi
43	170	15.7	234	6	AX244726	AX244726 Sequence
44	165.2	15.3	1177	10	RNP2V4REC	Y11433 R.norvegicu
45	164.8	15.2	984	6	AR109633	AR109633 Sequence

ALIGNMENTS

RESULT 1	AX458306	Sequence 1	1082 bp	DNA	linear	PAT 08-JUL-2002
AX458306	LOCUS	Sequence 1 from Patent EP1215214.				
DEFINITION	AX458306	AX458306				
ACCESSION	AX458306.1	GI:21725001				
VERSION	human.					
KEYWORDS	Homo sapiens					
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1					
AUTHORS	Fidock,M.D.					
TITLE	Novel polypeptide					
JOURNAL	Patent: EP 1215214-A 1 19-JUN-2002;					





Db 177 CTCCTGGAGGAGCCTCCCGGACATGAGAAAGTGGACATGAATACATCACAGGAACAA 236  
Qy 121 GGTCTCTGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCTGGCCTACAGTATC 180  
Db 237 GGTCTCTGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCTGGCCTACAGTATC 296  
Qy 181 ATCTTTATCTAGGCTGCCACTAAATAGCCACTGTCTGTGTGCACTCTCTGGGGCCAAACC 240  
Db 297 ATCTTTATCTAGGCTGCCACTAAATAGCCACTGTCTGTGTGCACTCTCTGGGGCCAAACC 356  
Qy 241 AAGCCTGGAGTGTGCCACCACTATCTGTGGAACCTGTAGTGGCGGCACTGCTTTAT 300  
Db 357 AAGCCTGGAGTGTGCCACCACTATCTGTGGAACCTGTAGTGGCGGCACTGCTTTAT 416  
Qy 301 GTGCTATTGGCCTCTCTCATCATCACTACTACCTACTAGTAGAGTGGCCTTCGGGGAG 360  
Db 417 GTGCTATTGGCCTCTCTCATCATCACTACTACCTACTAGTAGAGTGGCCTTCGGGGAG 476  
Qy 361 GTGCTCTGCAAGCTGGTGCACCTCTCTGTATATCAACCTTTACGGCAGCATCTGCTG 420  
Db 477 GTGCTCTGCAAGCTGGTGCACCTCTCTGTATATCAACCTTTACGGCAGCATCTGCTG 536  
Qy 421 CTGACCTGCATCTGTGCGACAGTTCCTAGTGTGTGCCACCACTGTGTTCGCTGCC 480  
Db 537 CTGACCTGCATCTGTGCGACAGTTCCTAGTGTGTGCCACCACTGTGTTCGCTGCC 596  
Qy 481 TACCGAGCCGAGGAGTCTGCTGGCTGGGACAGCACCACTGGCGCCCTGTGTGCTCTC 540  
Db 597 TACCGAGCCGAGGAGTCTGCTGGCTGGGACAGCACCACTGGCGCCCTGTGTGCTCTC 656  
Qy 541 CAGCTGTGCCCCACTGGCTCTCCACAGGACTACATCAATGGCAGATGATCTGG 600  
Db 657 CAGCTGTGCCCCACTGGCTCTCCACAGGACTACATCAATGGCAGATGATCTGG 716  
Qy 601 TATGACATGACGAGGAGAAATTTGATCGGCTTTTGCCTACGGCATAGTTCTGACA 660  
Db 717 TATGACATGACGAGGAGAAATTTGATCGGCTTTTGCCTACGGCATAGTTCTGACA 776  
Qy 661 TTGCTGGCTTTCTTCTCCCTCTGCTGCTATTTTGGTGTGCTATTCACTGATGCTCAGGAG 720  
Db 777 TTGCTGGCTTTCTTCTCCCTCTGCTGCTATTTTGGTGTGCTATTCACTGATGCTCAGGAG 836  
Qy 721 CCTGATCAAGCCAGAGGAGAACTCATGAGACAGGCAACACAGCCCGAGCCAGGTCCAT 780  
Db 837 CCTGATCAAGCCAGAGGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGTCCAT 896  
Qy 781 CCGGACCATCTACTGGTGTGTGGCCTCTTCAACCTCTGTTTGTGGCCTTCCATATCAC 840  
Db 897 CCGGACCATCTACTGGTGTGTGGCCTCTTCAACCTCTGTTTGTGGCCTTCCATATCAC 956  
Qy 841 TCGCTCTCTTACCTCAGCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTCATGAGC 900  
Db 957 TCGCTCTCTTACCTCAGCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTCATGAGC 1016  
Qy 901 AGCCAGTGTGGCTACAAGATATGAGGCCCTCTGGTGTGCTGTGAGCTGTGAGCTGCTCAACCC 960  
Db 1017 AGCCAGTGTGGCTACAAGATATGAGGCCCTCTGGTGTGCTGTGAGCTGTGAGCTGCTCAACCC 1076  
Qy 961 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAACTGAGG 1020  
Db 1077 AGTCTGTACTTTCTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAACTGAGG 1136  
Qy 1021 CAGAAACAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGCTTGAACAGATCT 1080  
Db 1137 CAGAAACAGTTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGCTTGAACAGATCT 1196  
Qy 1081 GG 1082  
Db 1197 GG 1198

AX365511 3143 bp DNA linear PAT 21-FEB-2002  
LOCUS Sequence 1 from Patent WO0185764.  
DEFINITION AX365511  
ACCESSION AX365511  
VERSION AX365511.1 GI:18873666  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1  
REFERENCE Ramakrishnan.S.  
AUTHORS Regulation of human p2y-like g protein-coupled receptor  
TITLE Patent: WO 0185764-A 1 15-NOV-2001;  
JOURNAL Bayer Aktiengesellschaft (DE)  
FEATURES Location/Qualifiers  
source  
1..3143  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
520..>1374  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD24312.1"  
/db\_xref="GI:18873667"  
/translation="MEKVDMMTSEQGLQCFSEKYQVYLSLAYSIIFILGPNLNGTV  
LWHSWGOTKRWSCATTVLNLWADLLVLLPFLITYSLDDRWPFGLKLLKLVHFLF  
YINLYSGTILLTCTISVHQFVCHPLCSLPYTRHAWLGTSTTVALVQLPTLAF  
SHTDYINGOMIWDYDNTSQENFDRLFAYGIVLTLSGFLSLGHFVFLFDGQEPDQARG  
EPHEDRQSPQVHPDHPGTGVWPLPLFCALPLHLLSFAFSLGALPDGSGQC  
LQDMEASGE"  
BASE COUNT 818 a 771 c 726 g 828 t  
ORIGIN  
Query Match 100.0%; Score 1082; DB 6; Length 3143;  
Best Local Similarity 100.0%; Pred. No. 1.7e-268;  
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGCTGTCCATTTTCTCTCCAGGGAAGCAAGCGGAGCGCTCGTGAGCTCTG 60  
Db 436 ATGCTGTCCATTTTCTCTCCAGGGAAGCAAGCGGAGCGCTCGTGAGCTCTG 495  
Qy 61 CTCCTGGAGGAGCCTCCGGGACATGGAGAAGTGGAGATGAATACATCACAGGAACAA 120  
Db 496 CTCCTGGAGGAGCCTCCGGGACATGGAGAAGTGGAGATGAATACATCACAGGAACAA 555  
Qy 121 GGTCTCTGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCTGGCCTACAGTATC 180  
Db 556 GGTCTCTGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCTGGCCTACAGTATC 615  
Qy 181 ATCTTTATCTAGGCTGCCACTAAATAGCCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 240  
Db 616 ATCTTTATCTAGGCTGCCACTAAATAGCCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 675  
Qy 241 AAGCCTGGAGTGTGCCACCACTATCTGTGGAACCTGTAGTGGCGGCACTGCTTTAT 300  
Db 676 AAGCCTGGAGTGTGCCACCACTATCTGTGGAACCTGTAGTGGCGGCACTGCTTTAT 735  
Qy 301 GTGCTATTGGCCTCTCTCATCATCACTACTACCTACTAGTAGAGTGGCCTTCGGGGAG 360  
Db 736 GTGCTATTGGCCTCTCTCATCATCACTACTACCTACTAGTAGAGTGGCCTTCGGGGAG 795  
Qy 361 GTGCTCTGCAAGCTGGTGCACCTCTGCTTTCTATATCAACCTTTACGGCAGCATCTGCTG 420  
Db 796 GTGCTCTGCAAGCTGGTGCACCTCTGCTTTCTATATCAACCTTTACGGCAGCATCTGCTG 855  
Qy 421 CTGACCTGCATCTGTGCGACAGTTCCTAGTGTGTGCCACCACTGTGTTCGCTGCC 480  
Db 856 CTGACCTGCATCTGTGCGACAGTTCCTAGTGTGTGCCACCACTGTGTTCGCTGCC 915  
Qy 481 TACCGAGCCGAGGAGTCTGCTGGCTGGGACAGCACCACTGGCGCCCTGTGTGCTCTC 540  
Db 916 TACCGAGCCGAGGAGTCTGCTGGCTGGGACAGCACCACTGGCGCCCTGTGTGCTCTC 975



Qy	421	CTGACCTGCATCTCTGTGCAACCAAGTTCCTAGGTGTGTGCCACCCACTGTGTTCGCTGCC	480
Db	123308	CTGACCTGCATCTCTGTGCAACCAAGTTCCTAGGTGTGTGCCACCCACTGTGTTCGCTGCC	123249
Qy	481	TACCGGACCCGCAAGCATGCTGTGGTGCGGACCAAGCAACCACTGGGCCCTGTGTGCTCTC	540
Db	123248	TACCGGACCCGCAAGCATGCTGTGGTGCGGACCAAGCAACCACTGGGCCCTGTGTGCTCTC	123189
Qy	541	CAGTCTGTGCCACACATGGGCTTCTCCACACGAGTACATCAATGGCCAGATGATCTGG	600
Db	123188	CAGTCTGTGCCACACATGGGCTTCTCCACACGAGTACATCAATGGCCAGATGATCTGG	123129
Qy	601	TATGACATGACCAAGCAAGAGAAATTTGATGGCTTTTTCCTACGGCATAGTCTTGACA	660
Db	123128	TATGACATGACCAAGCAAGAGAAATTTGATGGCTTTTTCCTACGGCATAGTCTTGACA	123069
Qy	661	TTGTCTGGCTTCTTTCCTCCCTTGGTCATTTTGGTGTGCTATTACTCATGGTCAGGAG	720
Db	123068	TTGTCTGGCTTCTTTCCTCCCTTGGTCATTTTGGTGTGCTATTACTCATGGTCAGGAG	123009
Qy	721	CCTGATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCGAGCCAGGTCCAT	780
Db	123008	CCTGATCAAGCCAGAGAGAACTCATGAGGACAGGCAACACAGCCGAGCCAGGTCCAT	122949
Qy	781	CCGGACCATCTACTGTGTGTGGCCCTTTCACCCCTCTGTTTGTGGCTTCCATATCAC	840
Db	122948	CCGGACCATCTACTGTGTGTGGCCCTTTCACCCCTCTGTTTGTGGCTTCCATATCAC	122899
Qy	841	TCGCTCTCTTACCTCACATCTGCTTTCTGCTTCTTCAGGACTGCCAGCTCTTGATGCG	900
Db	122888	TCGCTCTCTTACCTCACATCTGCTTTCTGCTTCTTCAGGACTGCCAGCTCTTGATGCG	122829
Qy	901	AGCCAGTGTGCGCTACAGATATGGAGGCTCTGTGCTGTGAGGAGTGCCTCAACCC	960
Db	122828	AGCCAGTGTGCGCTACAGATATGGAGGCTCTGTGCTGTGAGGAGTGCCTCAACCC	122769
Qy	961	AGTCCTCTACTTCTTTCAAGGGGGCAAAATAGAGTCAAGGCTCTCCAGAACTGAGG	1020
Db	122768	AGTCCTCTACTTCTTTCAAGGGGGCAAAATAGAGTCAAGGCTCTCCAGAACTGAGG	122709
Qy	1021	CAGAACAAATGGGTGAGCATCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT	1080
Db	122708	CAGAACAAATGGGTGAGCATCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCT	122649
Qy	1081	GG 1082	
Db	122648	GG 122647	
RESULT	5		
LOCUS	AC092999	167084 bp	DNA
DEFINITION	Homo sapiens 3 BAC RP11-170K4 (Roswell Park Cancer Institute Human BAC library) complete sequence.	linear	PRI 30-JUL-2001
ACCESSION	AC092999		
VERSION	AC092999.3	GI:19033407	
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 167084) Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayale,M., Banks,T., Barbra,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Bathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denny,A.L., Ding,Y., Dinh,H.H.,		

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabel, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharge, O., Lieu, C., Liu, J., Liu, W., Louised, H., Lozada, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metsker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Saverly, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 167084)  
Worley, K.C.  
Direct Submission  
Submitted (09-AUG-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 167084)  
Worley, K.C.  
Direct Submission  
Submitted (01-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 167084)  
Worley, K.C.  
Direct Submission  
Submitted (29-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
5 (bases 1 to 167084)  
Worley, K.C.  
Direct Submission  
Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Mar 1, 2002 this sequence version replaced gi:18139250.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

#### ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,

unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

#### QUALSTAT-REPORT

FEATURES	Source	Location/Qualifiers
		1..167084
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="3"
		/clone="RP11-170K4"
		/complement(1..2005)
		/note="overlaps bases 1..2005 of clone AC092997"
misc_feature		189..376
repeat_region		/rpt_family="MIR"
repeat_region		complement(997..1123)
repeat_region		/rpt_family="L2"
repeat_region		1999..2058
repeat_region		/rpt_family="L2"
repeat_region		2096..2404
repeat_region		/rpt_family="AluSx"
repeat_region		2422..2529
repeat_region		/rpt_family="L2"
repeat_region		2975..3004
repeat_region		/rpt_family="AT-rich"
repeat_region		3134..3163
repeat_region		/rpt_family="AT-rich"
repeat_region		4270..4290
repeat_region		/rpt_family="AT-rich"
repeat_region		complement(4421..4550)
repeat_region		/rpt_family="MIR"
repeat_region		5251..5642
repeat_region		/rpt_family="MLTK"
repeat_region		complement(5812..6366)
repeat_region		/rpt_family="MER34B"
repeat_region		6909..6941
repeat_region		/rpt_family="AT-rich"
repeat_region		7595..7717
repeat_region		/rpt_family="MER5A"
repeat_region		complement(7788..8198)
repeat_region		/rpt_family="MLTK"
repeat_region		complement(8536..8591)
repeat_region		/rpt_family="L2"
repeat_region		9634..9936
repeat_region		/rpt_family="AluSx"
repeat_region		complement(10766..10814)
repeat_region		/rpt_family="MIR"
repeat_region		complement(10848..10964)
repeat_region		/rpt_family="MIR"
repeat_region		17399..17431
repeat_region		/rpt_family="(TTAAA)n"
repeat_region		complement(17432..17546)
repeat_region		/rpt_family="MIR"
repeat_region		17637..18073



Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (20-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 10, 2000 this sequence version replaced gi:7248967.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www.seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
-----  
Center project name: L3302  
Center clone name: L2\_N17  
----- Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 152766 bases at least Q40  
Consensus quality: 158075 bases at least Q30  
Consensus quality: 160495 bases at least Q20  
Insert size: 174000; agarose-fp  
Quality coverage: 162558; sum-of-contigs  
Quality coverage: 4.5 in Q20 bases; agarose-fp  
Quality coverage: 4.9 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 15 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
-----

1 7254: contig of 7254 bp in length  
7255 7354: gap of 100 bp  
7355 8612: contig of 1258 bp in length  
8613 8712: gap of 100 bp  
8713 11463: contig of 2751 bp in length  
11464 11563: gap of 100 bp  
11564 14507: contig of 2944 bp in length  
14508 14607: gap of 100 bp  
14608 22452: contig of 7845 bp in length  
22453 22552: gap of 100 bp  
22553 35086: contig of 12534 bp in length  
35087 35186: gap of 100 bp  
35187 44278: contig of 9092 bp in length  
44279 44378: gap of 100 bp  
44379 56786: contig of 12408 bp in length  
56787 56886: gap of 100 bp  
56887 65270: contig of 8384 bp in length  
65271 65370: gap of 100 bp  
65371 76394: contig of 11024 bp in length  
76395 76494: gap of 100 bp  
76495 100320: contig of 23826 bp in length  
100321 100420: gap of 100 bp  
100421 113127: contig of 12707 bp in length  
113128 113227: gap of 100 bp  
113228 126756: contig of 13529 bp in length  
126757 126856: gap of 100 bp  
126857 151061: contig of 24205 bp in length  
151062 151161: gap of 100 bp  
151162 163958: contig of 12797 bp in length.  
Location/Qualifiers  
1. .163958  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-12N17"  
/clone\_lib="RPC1-11 Human Male BAC"

FEATURES  
source

misc\_feature  
1. 7254  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
7355. .8612  
misc\_feature  
/note="assembly\_fragment"  
8713. .11463  
misc\_feature  
/note="assembly\_fragment"  
11564. .14507  
misc\_feature  
/note="assembly\_fragment"  
14608. .22452  
misc\_feature  
/note="assembly\_fragment"  
22553. .35086  
misc\_feature  
/note="assembly\_fragment"  
35187. .44278  
misc\_feature  
/note="assembly\_fragment"  
44379. .56786  
misc\_feature  
/note="assembly\_fragment"  
56887. .65270  
misc\_feature  
/note="assembly\_fragment"  
65371. .76394  
misc\_feature  
/note="assembly\_fragment"  
76495. .100320  
misc\_feature  
/note="assembly\_fragment"  
100421. .113127  
misc\_feature  
/note="assembly\_fragment"  
113228. .126756  
misc\_feature  
/note="assembly\_fragment"  
126857. .151061  
misc\_feature  
/note="assembly\_fragment"  
151162. .163958  
/note="assembly\_fragment"  
clone\_end:T7  
vector\_side:right

BASE COUNT 49230 a 32369 c 32422 g 48536 t 1401 others  
ORIGIN

Query Match 99.0%; Score 1071; DB 2: Length 163958;  
Best Local Similarity 99.9%; Pred. No. 1.6e-265;  
Matches 1082; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 ATGCTGTCATTTTCCTCCAGGGGAGCAGAGCGGGAGCGCTGGAGCTCTG 60  
Db 98103 ATGCTGTCATTTTCCTCCAGGGGAGCAGAGCGGGAGCGCTGGAGCTCTG 98162  
Qy 61 CTCCTGGAGGAGCCTCCCGGGACATGGAGAGGTGGACATACATACAGAACAA 120  
Db 98163 CTCCTGGAGGAGCCTCCCGGGACATGGAGAGGTGGACATACATACAGAACAA 98222  
Qy 121 GGTCTCTGCCAGTTCTCAGAGAAAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 180  
Db 98223 GGTCTCTGCCAGTTCTCAGAGAAAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC 98282  
Qy 181 ATCTTTATCTAGGCTGCCACTAAATGGCCTGCTGTGGCCTCTCTGGGCGCAACCC 240  
Db 98283 ATCTTTATCTAGGCTGCCACTAAATGGCCTGCTGTGGCCTCTCTGGGCGCAACCC 98342  
Qy 241 AAGCGCTGGAGCTGTGCCACCACCTATCTGTGAACCTGTGTGGCGGAGCTGCTTTAT 300  
Db 98343 AAGCGCTGGAGCTGTGCCACCACCTATCTGTGAACCTGTGTGGCGGAGCTGCTTTAT 98402  
Qy 301 GTGTAATGGCTTCTCCTCATCATCACCTACTACCTAGATACAGGTGGCCCTTCGGGGAG 360  
Db 98403 GTGTAATGGCTTCTCCTCATCATCACCTACTACCTAGATACAGGTGGCCCTTCGGGGAG 98462  
Qy 361 CTGCTCTGCAAGCTGGTGCACCTCTCTGTCTATATACCTTTTACGGCAGCATCTGCTG 420  
Db 98463 CTGCTCTGCAAGCTGGTGCACCTCTCTGTCTATATACCTTTTACGGCAGCATCTGCTG 98522  
Qy 421 CTGACCTGCATCTCTGTGCACAGTTCTCTAGGTGTGTGCCACCCACTGTGTTCGCTGCC 480  
Db 98523 CTGACCTGCATCTCTGTGCACAGTTCTCTAGGTGTGTGCCACCCACTGTGTTCGCTGCC 98582

QY 481 TACCGACCGCAGGCGATGCTGGCTGGGACACAGCACACCTGGGCCCTGGTGGCTCTC 540  
Db 98583 TACCGACCGCAGGCGATGCTGGCTGGGACACAGCACACCTGGGCCCTGGTGGCTCTC 98642  
QY 541 CAGCTGCTGCCACACTGGCTTCTCCACACAGGACTACATCAATGGCCAGATGATCTGG 600  
Db 98643 CAGCTGCTGCCACACTGGCTTCTCCACACAGGACTACATCAATGGCCAGATGATCTGG 98702  
QY 601 TATGACATGACACGAGCAAGAAATTTGATCGGCTTTTGGCCCTACGSCATAGTTCTGACA 660  
Db 98703 TATGACATGACACGAGCAAGAAATTTGATCGGCTTTTGGCCCTACGSCATAGTTCTGACA 98762  
QY 661 TTGCTGGCTTTCTTTCCCTCCTTGCTCAATTTGGTGTGCTATTTCACTGATGCTCAGGAG 720  
Db 98763 TTGCTGGCTTTCTTTCCCTCCTTGCTCAATTTGGTGTGCTATTTCACTGATGCTCAGGAG 98822  
QY 721 CTGATCAAGCCAGAGGAACTCATGAGGACAGCAACAGCCCGAGCCAGGTCAT 780  
Db 98823 CTGATCAAGCCAGAGGAACTCATGAGGACAGCAACAGCCCGAGCCAGGTCAT 98882  
QY 781 CCGGACCATCTACTGGTGTGGCCCTCTTACCCCTCTGTTTGTGCCCTTCCATATCAC 840  
Db 98883 CCGGACCATCTACTGGTGTGGCCCTCTTACCCCTCTGTTTGTGCCCTTCCATATCAC 98942  
QY 841 TCCTCCTTTACTCCTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC 900  
Db 98943 TCCTCCTTTACTCCTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC 99002  
QY 901 AG-CCAGTGTGGCTACAGATATGAGGCTCTGTGTGCTGCTTCCATATCAC 959  
Db 99003 AGCCAGTGTGGCTACAGATATGAGGCTCTGTGTGCTGCTTCCATATCAC 99062  
QY 960 CAGTCTGTACTTCTTCAAGGGGGCAAAATAGATCAGGCTCCTCCAGAACTGAG 1019  
Db 99063 CAGTCTGTACTTCTTCAAGGGGGCAAAATAGATCAGGCTCCTCCAGAACTGAG 99122  
QY 1020 GCAGACAAAGTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGTTGAACAGATC 1079  
Db 99123 GCAGACAAAGTGGGTGAGCATCCAGCTGGGAGGAGAGATGCCAGGTTGAACAGATC 99182  
QY 1080 TGG 1082  
Db 99183 TGG 99185

RESULT 7  
AF411108 1019 bp DNA linear PRI 01-NOV-2001  
LOCUS Homo sapiens GPR79 pseudogene, complete sequence.  
DEFINITION AF411108  
ACCESSION AF411108.1 GI:16566321  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1019)  
AUTHORS Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,  
Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.  
Discovery and mapping of ten novel G protein-coupled receptor genes  
Gene 275 (1), 83-91 (2001).  
JOURNAL 11574155  
MEDLINE 21458557  
PUBMED  
REFERENCE 2 (bases 1 to 1019)  
AUTHORS Lee,D.K., Nguyen,T., Lynch,K.R., Cheng,R., Vanti,W.B., Arkhitko,O.,  
Lewis,T., Evans,J.F., George,S.R. and O'Dowd,B.F.  
Direct Submission  
Submitted (17-AUG-2001) Department of Pharmacology, University of  
Toronto, 8 Taddle Creek Rd. Rm. 4353, Toronto, Ont M5S 1A8, Canada  
JOURNAL  
TITLE Location/Qualifiers  
FEATURES  
source 1. .1019  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

gene 1. .1019  
/gene="GPR79"  
/pseudo  
BASE COUNT 200 a 308 c 252 g 259 t  
ORIGIN  
Query Match 93.1%; Score 1007; DB 9; Length 1019;  
Best Local Similarity 99.9%; Pred. No. 3.6e-249;  
Matches 1018; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1 ATGCTCTCCATTTTGTCTTCTCCAGGGGAAGAGGGAGCGTGTGTGAGCTCTG 60  
Db 1 ATGCTCTCCATTTTGTCTTCTCCAGGGGAAGAGGGAGCGTGTGTGAGCTCTG 60  
QY 61 CTCCTGAGGGAGCCCTCCCGGACATGGAGAGGTGGAGATGAATACATCACAGGAACA 120  
Db 61 CTCCTGAGGGAGCCCTCCCGGACATGGAGAGGTGGAGATGAATACATCACAGGAACA 120  
QY 121 GGTCTCTGCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGGCTTACAGTATC 180  
Db 121 GGTCTCTGCAGTTCTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGGCTTACAGTATC 180  
QY 181 ATCTTTTATCCTAGGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 240  
Db 181 ATCTTTTATCCTAGGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC 240  
QY 241 AAGCGCTGGAGCTGCCACCACCTATCTGGTGAACCTGATGTGGCGGACCTGCTTTAT 300  
Db 241 AAGCGCTGGAGCTGCCACCACCTATCTGGTGAACCTGATGTGGCGGACCTGCTTTAT 300  
QY 301 GTGCTATTGCGCTTCTCATCATCATCTACTAGTACAGAGTGGGCTTTCGGGGAG 360  
Db 301 GTGCTATTGCGCTTCTCATCATCATCTACTAGTACAGAGTGGGCTTTCGGGGAG 360  
QY 361 CTGCTCTGAAGCTGGTGACATCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCTG 420  
Db 361 CTGCTCTGAAGCTGGTGACATCTCTTCTATATCAACCTTTACGGCAGCATCTCTGCTG 420  
QY 421 CTGACCTGCTATCTCTGACACAGTTCCTAGGTGTGTGCCACCACTGTCTTCGCTGCC 480  
Db 421 CTGACCTGCTATCTCTGACACAGTTCCTAGGTGTGTGCCACCACTGTCTTCGCTGCC 480  
QY 481 TACCGACCGCAGGCGATGCTGGTGGGACCAAGCAACCTGCTGGGCTTGGTGGTCTC 540  
Db 481 TACCGACCGCAGGCGATGCTGGTGGGACCAAGCAACCTGCTGGGCTTGGTGGTCTC 540  
QY 541 CAGCTGCTGCCACACTGGCTTCTCCACACGACTACATCAATGCCAGATGATCTGG 600  
Db 541 CAGCTGCTGCCACACTGGCTTCTCCACACGACTACATCAATGCCAGATGATCTGG 600  
QY 601 TATGACATGACACGAGCAAGAAATTTGATCGGCTTTTGGCTTACGSCATAGTTCTGACA 660  
Db 601 TATGACATGACACGAGCAAGAAATTTGATCGGCTTTTGGCTTACGSCATAGTTCTGACA 660  
QY 661 TTGCTGGCTTTCTTTCCCTCCTTGCTCAATTTGGTGTGCTATTTCACTGATGCTCAGGAG 720  
Db 661 TTGCTGGCTTTCTTTCCCTCCTTGCTCAATTTGGTGTGCTATTTCACTGATGCTCAGGAG 720  
QY 721 CCTGATCAAGCCAGAGGAACTCATGAGGACAGCAACAGCCCGAGCCAGGTCAT 780  
Db 721 CCTGATCAAGCCAGAGGAACTCATGAGGACAGCAACAGCCCGAGCCAGGTCAT 780  
QY 781 CCGGACCATCTACTGGTGTGGCCCTCTTACCCCTCTGTTTGTGCCCTTCCATATCAC 840  
Db 781 CCGGACCATCTACTGGTGTGGCCCTCTTACCCCTCTGTTTGTGCCCTTCCATATCAC 840  
QY 841 TCCTCCTTTACTCCTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC 900  
Db 841 TCCTCCTTTACTCCTCACCATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGC 900  
QY 901 AG-CCAGTGTGGCTACAGATATGAGGCGCTCTGGTGTGCTGTGAGCAGCTGCTCAACC 959  
Db 901 AGCCAGTGTGGCTACAGATATGAGGCGCTCTGGTGTGCTGTGAGCAGCTGCTCAACC 960

```
Qy 960 CAGTCCTGTACTTTCTTCAAGGGGGCAAAATAGAGTCAGGCTCCTCCAGAACTGA 1018
|||||
Db 961 CAGTCTGTACTTTCTTCAAGGGGGCAAAATAGAGTCAGGCTCCTCCAGAACTGA 1019

RESULT 8
AX458308
LOCUS AX458308 1020 bp DNA linear PAT 08-JUL-2002
DEFINITION Sequence 3 from Patent EP1215214.
ACCESSION AX458308
VERSION AX458308.1 GI:21725002
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Fidock, M.D.
TITLE Novel polypeptide
JOURNAL Patent: EP 1215214-A 3 19-JUN-2002;
Pfizer Limited (GB); PFIZER INC. (US)
FEATURES
LOCATION/Qualifiers
source 1..1020
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 200 a 309 c 252 g 259 t
ORIGIN
Query Match 92.7%; Score 1002.8; DB 6; Length 1020;
Best Local Similarity 99.6%; Pred. No. 4.4e-248;
Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 1 ATGCTGTCCATTTTGTCTTCCAGGGGAAGCAGCGGAGCGCTGTGGAGCTCTG 60
|||||
Db 1 ATGCTGTCCATTTTGTCTTCCAGGGGAAGCAGCGGAGCGCTGTGGAGCTCTG 60

Qy 61 CTCCTGGAGGAGCTCCCGGGACATGGAGAAGTGGACATGAATACATCAGAGACAA 120
|||||
Db 61 CTCCTGGAGGAGCTCCCGGGACATGGAGAAGTGGACATGAATACATCAGAGACAA 120

Qy 121 GCTCTGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCTACAGTATC 180
|||||
Db 121 GCTCTGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCTACAGTATC 180

Qy 181 ATCTTTATCTTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCAAAAC 240
|||||
Db 181 ATCTTTATCTTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCAAAAC 240

Qy 241 AAGCGCTGGAGCTGGCACCACCTATCTGGTGAACCTGATGGTGGCGGACCTGCTTTAT 300
|||||
Db 241 AAGCGCTGGAGCTGGCACCACCTATCTGGTGAACCTGATGGTGGCGGACCTGCTTTAT 300

Qy 301 GTGCTATTGCCCTTCTCATCATCATCACTACTACTAGATGACAGTGGCCCTTCGGGGAG 360
|||||
Db 301 GTGCTATTGCCCTTCTCATCATCATCACTACTACTAGATGACAGTGGCCCTTCGGGGAG 360

Qy 361 CTGCTCTGCAAGCTGGTGCAGTCTTCTGTATATCAACCTTTACGGCAGCATCTGCTG 420
|||||
Db 361 CTGCTCTGCAAGCTGGTGCAGTCTTCTGTATATCAACCTTTACGGCAGCATCTGCTG 420

Qy 421 CTGACCTGCATCTCTGTGCACCACTTCTGTAGGTGTGGCCACCACTGTGTTCGCGCC 480
|||||
Db 421 CTGACCTGCATCTCTGTGCACCACTTCTGTAGGTGTGGCCACCACTGTGTTCGCGCC 480

Qy 481 TACCGGACCCGACGATGCTGTGCTGGCACACAGCACCACCTGGCGCTGTGGTGCCTC 540
|||||
Db 481 TACCGGACCCGACGATGCTGTGCTGGCACACAGCACCACCTGGCGCTGTGGTGCCTC 540

Qy 541 CAGCTGTGCCACACGAGGCTTCTCCACACGAGCTACATCAATGCGCAGATGATCTGG 600
|||||
Db 541 CAGCTGTGCCACACGAGGCTTCTCCACACGAGCTACATCAATGCGCAGATGATCTGG 600
```

```
Qy 601 TATGACATGACCAGCCAGAGAAATTTTGATCGGCTTTTGTGCTACGGCATAGTCTTGACA 660
|||||
Db 601 TATGACATGACCAGCCAGAGAAATTTTGATCGGCTTTTGTGCTACGGCATAGTCTTGACA 660

Qy 661 TTGTCTGGCTTTCTTT--CCCTCCTTGTCTATTTTGGTGTCTATTCTACTGATGGTCAGG 718
|||||
Db 661 TTGTCTGGCTTTCTTTTCCCTCCTTGTCTATTTTGGTGTCTATTCTACTGATGGTCAGG 720

Qy 719 AGCTGTATCAAGCCAGAGGAACCTCATGAGGACAGCAACACAGCCCGAGCCAGGTCC 778
|||||
Db 721 AGCTGTATCAAGCCAGAGGAACCTCATGAGGACAGCAACACAGCCCGAGCCAGGTCC 780

Qy 779 ATCCGGACCATCTACTGCTGTGGCTCTTCAACCTCTCTTTTGTGCCCTTCCATATC 838
|||||
Db 781 ATCCGGACCATCTACTGCTGTGGCTCTTCAACCTCTCTTTTGTGCCCTTCCATATC 840

Qy 839 ACTCGCTCTTCTACTCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATG 898
|||||
Db 841 ACTCGCTCTTCTACTCACCATCTGCTTCTGCTTCTCAGGACTGCCAGCTCTTGATG 900

Qy 899 GCAGCCAGTGTGGCTTACAAGATATGGAGGCTCTGCTGAGTGTGAGCAGCTGCCTCAAC 958
|||||
Db 901 GCAGCCAGTGTGGCTTACAAGATATGGAGGCTCTGCTGAGTGTGAGCAGCTGCCTCAAC 960

Qy 959 CCAGTCTGTACTTTTCTTCAAGGGGGCAAAATAGAGTCAGGCTCCTCCAGAACTGA 1018
|||||
Db 961 CCAGTCTGTACTTTTCTTCAAGGGGGCAAAATAGAGTCAGGCTCCTCCAGAACTGA 1020

RESULT 9
AX375230
LOCUS AX375230 1076 bp DNA linear PAT 01-MAR-2002
DEFINITION Sequence 20 from Patent WO0210387.
ACCESSION AX375230
VERSION AX375230.1 GI:19169948
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R.,
Elliot, V.S., Rankumar, J., Baughn, M.R., Kallick, D.A., Walia, N.K.,
Hafalia, A.J., Yao, M.G., Lu, Y., Tribouley, C.M., Pollicky, J.L.,
Kearney, L., Graul, R.C., Warren, B.A. and Ding, L.
TITLE G-protein coupled receptors
JOURNAL Patent: WO 0210387-A 20 07-FEB-2002;
Incyte Genomics, Inc. (US)
FEATURES
LOCATION/Qualifiers
source 1..1076
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 7474806CB1"
BASE COUNT 211 a 318 c 274 g 273 t
ORIGIN
Query Match 92.7%; Score 1002.8; DB 6; Length 1076;
Best Local Similarity 99.6%; Pred. No. 4.4e-248;
Matches 1016; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

Qy 1 ATGCTGTCCATTTTGTCTTCCAGGGGAAGCAGAGCGGAGCGCTCGTGGAGCTCTG 60
|||||
Db 57 ATGCTGTCCATTTTGTCTTCCAGGGGAAGCAGAGCGGAGCGCTCGTGGAGCTCTG 116

Qy 61 CTCCTGGAGGAGCTCCCGGGACATGGAGAAGTGGACATGAATACATCAGAGACAA 120
|||||
Db 117 CTCCTGGAGGAGCTCCCGGGACATGGAGAAGTGGACATGAATACATCAGAGACAA 176

Qy 121 GGTCTCTGCCAGTCTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCTACAGTATC 180
|||||
Db 177 GGTCTCTGCCAGTCTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCTACAGTATC 236

Qy 181 ATCTTTATCTTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCAAAAC 240
```



|||||  
Db 237 ATCTTTATCTAGGCTGCGCACTAAATGCACTGTCTTGCGCACTCTCGGGCCAAACC 296  
|||  
QY 241 AAGCGCTGAGCTGTGCCACCACTATCTGTGAACCTCATGTGTGCCACCTGCTTTAT 300  
|||||  
Db 297 AAGCGCTGAGCTGTGCCACCACTATCTGTGAACCTCATGTGTGCCACCTGCTTTAT 356  
|||||  
QY 301 GTGCTATTGCCCTTCTCATCATCACTACTCACTAGATGACAGGTGGCCCTTCGGGGAG 360  
|||||  
Db 357 GTGCTATTGCCCTTCTCATCATCACTACTCACTAGATGACAGGTGGCCCTTCGGGGAG 416  
|||||  
QY 361 GTGCTCTGAAGTGTGACCTCTCTGTTCTATATCAACCTTTACGGCAGCATCTGCTG 420  
|||||  
Db 417 CTGCTCTGAAGTGTGACCTCTCTGTTCTATATCAACCTTTACGGCAGCATCTGCTG 476  
|||||  
QY 421 CTGACCTGATCTCTGTGACCAAGTTCCTAGTGTGTGCGCACCCACTGTGTCGCTGCC 480  
|||||  
Db 477 CTGACCTGATCTCTGTGACCAAGTTCCTAGTGTGTGCGCACCCACTGTGTCGCTGCC 536  
|||||  
QY 481 TACCGGACCCGAGGATGCTGCTGGGACCAAGACCACTGGCCCTGTGTGCTCTC 540  
|||||  
Db 537 TACCGGACCCGAGGATGCTGCTGGGACCAAGACCACTGGCCCTGTGTGCTCTC 596  
|||||  
QY 541 CAGCTGCTGCCACACTGCTCTCCACAGGACTATCAATGGCCAGATGATCTG 600  
|||||  
Db 597 CAGCTGCTGCCACACTGCTCTCCACAGGACTATCAATGGCCAGATGATCTG 656  
|||||  
QY 601 TATGACATGACCAAGCAAGAAATTTGATCGCTTTTGGCTTTTGGCTTTTGGCTTTTGG 660  
|||||  
Db 657 TATGACATGACCAAGCAAGAAATTTGATCGCTTTTGGCTTTTGGCTTTTGGCTTTTGG 716  
|||||  
QY 661 TTGCTGCTGCTCTCTTTT - CCCCTCTGCTCATTTTGGTGTGCTTTTGGCTTTTGGCTTTTGG 718  
|||||  
Db 717 TTGCTGCTGCTCTCTTTTCCCTCTCTTGGTCAATTTTGGTGTGCTTTTGGCTTTTGGCTTTTGG 776  
|||||  
QY 719 AGCTGATCAAGCCAGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGCTC 778  
|||||  
Db 777 AGCTGATCAAGCCAGAGAACTCATGAGGACAGGCAACACAGCCCGAGCCAGGCTC 836  
|||||  
QY 779 ATCCGACCATCTACTGTGTGCTGCTCTTCCACCTCTGTTTGGCTTTTGGCTTTTGGCTTTTGG 838  
|||||  
Db 837 ATCCGACCATCTACTGTGTGCTGCTCTTCCACCTCTGTTTGGCTTTTGGCTTTTGGCTTTTGG 896  
|||||  
QY 839 ACTGCTCTCTTACCTACCACTGCTCTGCTCTGCTTTCTCAGGACTGCCAGCTCTTTGATG 898  
|||||  
Db 897 ACTGCTCTCTTACCTACCACTGCTCTGCTCTGCTTTCTCAGGACTGCCAGCTCTTTGATG 956  
|||||  
QY 899 GCAGCCAGTGTGCCCTACAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCCCTCAAC 958  
|||||  
Db 957 GCAGCCAGTGTGCCCTACAGATATGAGGCTCTGCTGAGTGTGAGCAGCTGCCCTCAAC 1016  
|||||  
QY 959 CCAGTCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAAACTGA 1018  
|||||  
Db 1017 CCAGTCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAAGCTCTCCAGAAACTGA 1076  
|||||

RESULT 10  
AX365514  
LOCUS AX365514 850 bp DNA linear PAT 15-FEB-2002  
DEFINITION Sequence 4 from Patent WO0185764.  
ACCESSION AX365514  
VERSION AX365514.1 GI:18697083

KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Ramakrishnan,S.  
TITLE Regulation of human p2y-like g protein-coupled receptor  
JOURNAL Patent: WO 0185764-A 4 15-NOV-2001;  
Bayer Aktiengesellschaft (DE)  
Location/Qualifiers

RESULT 11  
AC121579

source 1..850  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 165 a 262 c 201 g 222 t  
ORIGIN  
Query Match 78.6%; Score 850; DB 6; Length 850;  
Best Local Similarity 100.0%; Pred. No. 1.3e-208;  
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 85 ATGGAGAAGGTGGACATGAATACATCACAGGAACAAGGTCTCTGCCAGTCTCTCAGAGAAG 144  
|||||  
Db 1 ATGGAGAAGGTGGACATGAATACATCACAGGAACAAGGTCTCTGCCAGTCTCTCAGAGAAG 60  
|||||  
QY 145 TACAACAAGTCTACCTCTCCCTGCGCTACAGTATCATCTTTATCTTAGGGCTGCCACTA 204  
|||||  
Db 61 TACAACAAGTCTACCTCTCCCTGCGCTACAGTATCATCTTTATCTTAGGGCTGCCACTA 120  
|||||  
QY 205 AATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACCACC 264  
|||||  
Db 121 AATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACCACC 180  
|||||  
QY 265 TATCTGGTGAACCTGATGTGGCGGACCTGCTTTATGTGCTATTGGCCCTTCCCTCATCATC 324  
|||||  
Db 181 TATCTGGTGAACCTGATGTGGCGGACCTGCTTTATGTGCTATTGGCCCTTCCCTCATCATC 240  
|||||  
QY 325 ACCTACTCACTAGATGACAGGTGGCCCTTTCGGGGAGCTGCTCTGCAAGCTGTGTGCACTTC 384  
|||||  
Db 241 ACCTACTCACTAGATGACAGGTGGCCCTTTCGGGGAGCTGCTCTGCAAGCTGTGTGCACTTC 300  
|||||  
QY 385 CTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGTCATCTGTGTGCACTTC 444  
|||||  
Db 301 CTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGTCATCTGTGTGCACTTC 360  
|||||  
QY 445 TTCTAGTGTGTGCGCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504  
|||||  
Db 361 TTCTAGTGTGTGCGCACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
|||||  
QY 505 CTGGGACCAAGCAGCACCACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 564  
|||||  
Db 421 CTGGGACCAAGCAGCACCACCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
|||||  
QY 565 TCCCAACAGCAGCTACATCAATAGGCCAGATGATCTGATGATGATGATGATGATGATGATGATGATG 624  
|||||  
Db 481 TCCCAACAGCAGCTACATCAATAGGCCAGATGATCTGATGATGATGATGATGATGATGATGATGATG 540  
|||||  
QY 625 TTTGATCGGCTTTTGGCTTACGCGCATAGTCTGACATTTGCTGGCTTTCTTTCCCTCCCTT 684  
|||||  
Db 541 TTTGATCGGCTTTTGGCTTACGCGCATAGTCTGACATTTGCTGGCTTTCTTTCCCTCCCTT 600  
|||||  
QY 685 GGTCATTTTGGTGTGCTATTCTACTGATGTCAGGAGCTGATCAAGCCAGAGGAGAACCT 744  
|||||  
Db 601 GGTCATTTTGGTGTGCTATTCTACTGATGTCAGGAGCTGATCAAGCCAGAGGAGAACCT 660  
|||||  
QY 745 CATGAGGACAGGCAACACAGCCCGAGCGGATCCATCCGAGCAGCTGCTGCTGCTGCTGCTGCTGCTG 804  
|||||  
Db 661 CATGAGGACAGGCAACACAGCCCGAGCGGATCCATCCGAGCAGCTGCTGCTGCTGCTGCTGCTGCTG 720  
|||||  
QY 805 CCTCTTCCACCTCTGTTTGGCTTCCATATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 864  
|||||  
Db 721 CCTCTTCCACCTCTGTTTGGCTTCCATATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
|||||  
QY 865 CTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGAGCCAGCTGTGGCTTACAGATATG 924  
|||||  
Db 781 CTTTCTGCTTTCTCAGGACTGCCAGCTCTTGATGGAGCCAGCTGTGGCTTACAGATATG 840  
|||||  
QY 925 GAGGCTCTG 934  
|||||  
Db 841 GAGGCTCTG 850  
|||||

LOCUS AC121579 209273 bp DNA linear HTG 20-MAY-2002  
DEFINITION Mus musculus chromosome UNK clone RP23-255H8, WORKING DRAFT  
SEQUENCE, 9 unordered pieces.  
AC121579  
AC121579  
AC121579.1 GI:20986651  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 209273)  
McPherson, J.D. and Waterston, R.H.  
JOURNAL The sequence of Mus musculus clone  
Unpublished  
REFERENCE 2 (bases 1 to 209273)  
McPherson, J.D. and Waterston, R.H.  
JOURNAL Direct Submission  
TITLE Submitted (20-MAY-2002) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
COMMENT

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
Contact: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
----- Project Information -----  
Center project name: M\_BA0255H8  
----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Chemistry: Dye-primer ET; 0% of reads  
Assembly: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 205908 bases at least Q40  
Consensus quality: 206285 bases at least Q30  
Consensus quality: 206424 bases at least Q20  
Insert size: 177000; agarose-fp  
Insert size: 208741; sum-of-contigs  
Quality coverage: 15.45 in Q20 bases; agarose-fp  
Quality coverage: 12.20 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 9 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1957: contig of 1957 bp in length  
\* 1958 2057: gap of unknown length  
\* 2058 7480: contig of 5423 bp in length  
\* 7481 7580: gap of unknown length  
\* 7581 21224: contig of 13644 bp in length  
\* 21225 21324: gap of unknown length  
\* 21325 40793: contig of 19469 bp in length  
\* 40794 40893: gap of unknown length  
\* 40894 73232: contig of 32339 bp in length  
\* 73233 73332: gap of unknown length  
\* 73333 113562: contig of 40230 bp in length  
\* 113563 113662: gap of unknown length  
\* 113663 163371: contig of 51709 bp in length  
\* 163372 165471: gap of unknown length  
\* 165472 208277: contig of 42806 bp in length  
\* 208278 208378: gap of unknown length  
\* 208379 209273: contig of 896 bp in length.

FEATURES  
Location/Qualifiers  
1..209273  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="UNK"  
/clone="RP23-255H8"

misc\_feature 1..1957  
/note="assembly\_name:Contig18"  
misc\_feature 2058..7480  
/note="assembly\_name:Contig19"  
misc\_feature 7581..21224  
/note="assembly\_name:Contig20"  
misc\_feature 21325..40793  
/note="assembly\_name:Contig21"  
misc\_feature 40894..73232  
/note="assembly\_name:Contig22"  
misc\_feature 73333..113562  
/note="assembly\_name:Contig23"  
misc\_feature 113663..165371  
/note="assembly\_name:Contig24"  
misc\_feature 165472..208277  
/note="assembly\_name:Contig25"  
misc\_feature 208378..209273  
/note="assembly\_name:Contig16"  
BASE COUNT 59147 a 42827 c 43921 g 62577 t 801 others  
ORIGIN  
Query Match 60.0%; Score 649.4; DB 2; Length 209273;  
Best Local Similarity 78.3%; Pred. No. 1.3e-156;  
Matches 837; Conservative 0; Mismatches 211; Indels 21; Gaps 4;  
Qy 1 ATGCTGTCATTGCTTCTCCAGGGGAAGCAGAGCGGAGCCGCTCGTGGAGCTCTG 60  
Db 37890 ATGCTGTCGCTTTATTTCTTCAGGAGAAAGCAGAGGTCGACGCC----- 37935  
Qy 61 CTCCTGGAGGAGCCTCCCGGACATGGAGAGTGGGACATACATACATACACAGCAAA 120  
Db 37936 ATCTTCGAGGAAGCCTTTTAGGATGGAGAGCTAGATATGAATGCTCAAAGGTCGT 37995  
Qy 121 GGTCTCTGCCAGTCTCAGAGAAGTACAAAGTCTACCTCTCCCTGGCCCTACAGTATC 180  
Db 37996 GCGCCTGCCACTTCTCAGAGAGCTACAAAGCTACCTATCATTAACCTACAGTCTC 38055  
Qy 181 ATCTTTATCTTAGGCTGCCACTAAATGGCAGCTGCTTGTGGCAGCTCTCGGGCCAAACC 240  
Db 38056 ATCTTTATCTTAGGATGCCCTCAATGGGCTGCTTTATGGCTCTCCTGTGGCCAAACC 38115  
Qy 241 AAGCGCTGAGCTGGCCACCACCTATCTGTGAACCTGATGGTGGCGGACCTGCTTTAT 300  
Db 38116 AAGCGCTGAGCTGGCCACCACCTATCTGTGAACCTGATGAACCTGGTGGCGGACCTGCTTTAT 38175  
Qy 301 GTGCTA---TTGGCCTTCTCTCATCATCACTACTACTAGATGACAGTGGCCCTTCGGG 357  
Db 38176 ATGCTAAGCCTGCCCTTCTCATCATCACTACCTACGCGGTGGTGACACCTGGCTCTTTGGG 38235  
Qy 358 GAGCTGCTGCAAGCTGGTGCACCTTCTCTATATCAACCTTTACGGCAGCATCTCTG 417  
Db 38236 GAGATGCTGTGAGGTTGGTGGGTTCTCTTCTACCAACCTCTACAGCAGCATCTCTG 38295  
Qy 418 CTGCTGACCTGCATCTCTGTGCACCACTTCTAGTGTGTGGCCACCTACTGTTTCGCTG 477  
Db 38296 CTGCTGACCTGCATCTCTGTGCACCACTTCTAGTGTGTGGCCACCTACTGTTTCGCTG 38355  
Qy 478 CCTTACCGGACCGGAGCATGCGCTGGTGGGACACACACCTGGGCCCTGTGTGTC 537  
Db 38356 TCCTACCGGACTCACCGGACCGCTGGCAGCAGCTGTGTCACCTGGACCTGTGTAGTT 38415  
Qy 538 CTCAGCTGCTGCCACACTGGCCTTCTCCACAGGACTACATCAATGGCCAGATGATC 597  
Db 38416 CTCCTAATTTGCTGCCGACTCTGTGCTTCTCCACACAGACTATGAATAAATGGCCAGGTGATC 38475  
Qy 598 TGGTATGACATGACCAAGAGAAATTTGATGGCTTTTGTGCTTGTGCTAGCGATGCTCTG 657  
Db 38476 TGGTATGACATGACCAAGCAGCGGAGCAGTGTGATCATTTCTTCGCTACAGCATAGTCTG 38535  
Qy 658 ACATGCTGCTGCTT--TCTTTCCCTCCTTGGTCAATTTGGTGTGCTATCACTGATGTC 715  
Db 38536 ACCTTATCTGGGTTGCTTCTCCCTCCTTGTATCATTTCTTGGTGTGCTACTCTCTGATGCTT 38595

```

QY 716 AGAGGCTGATCAAGCCAGGAGGAACCTCATAGAGCAGGCAACACAGCCCGAGCCAGG 775
Db 38596 AGGAGCTGACAAATCCAGTGGAGGACCTCCCAACAGCAGGCAACAGCCCGTGCAGG 38655
QY 776 TCCATCCGAGCATTCTACTGTGTGTGCTTTCACACCTCTGTGTTTGTGCGCTTCAT 835
Db 38656 TCTATTCGAGCATTCTGCTGTGTGTGCTTTCACACTCTGTGTTTGTGCGCTTCAC 38715
QY 836 ATCACTCGCTCTTACCTCACCATCTGCTTCTCTCTCAGGACTGCCAGCTCTTTG 895
Db 38716 ATTGCCGATTCTTCTACTTGATCCGCTTCTTGAAGTCACAGGACTGCCAGCTTTG 38775
QY 896 ATGGCAGCAGGTGCGCTTACAGATATGGAGCCCTCTGCTGAGTGTGAGCAGCTGCCTC 955
Db 38776 TGGGCAGCAGCATGCGCTTACAGGTCTGTAGGCTCTGTGTAAGCATGAGCAGCTGCCTC 38835
QY 956 AACCCAGTCTGCTACTTCTTTCAGGGGGCAAAATAGATCAGGCTCCTCCAGAAC 1015
Db 38836 ATCCAGTCTGCTACTTCTTGTGCA--CAGGGCAACACAGAGCCCAACTCTCCAGAGGC 38893
QY 1016 TGAGGCAGAACAGTGTGGGTGAGCATCCAGCTGGGAGGAGAGATGCC 1064
Db 38894 TGAGACAAACAGGTGGGTGAGCACCCTGCTGTGGAGGAGGACCC 38942

RESULT 12
AC116265/c
LOCUS
DEFINITION
AC116265
VERSION
AC116265.2 GI:21745788
KEYWORDS
HTG: HTGS-PHASE1.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 91107)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowle,S., Brivea,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burrett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.K., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louiseged,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenko,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Petters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 91107)
Worley,K.C.
Direct Submission
Submitted (26-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 91107)
Worley,K.C.
Direct Submission
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:19718459.
----- Genome Center -----
Center: Baylor College Of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: CTRS
Center clone name: CH230-18307
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 36430 bases at least Q40
Consensus quality: 38793 bases at least Q30
Consensus quality: 40859 bases at least Q20
-----

```

```

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 46 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1061: contig of 1061 bp in length
* 1062 1161: gap of unknown length
* 1162 2717: contig of 1556 bp in length
* 2718 2817: gap of unknown length
* 2818 4467: contig of 1650 bp in length
* 4468 4567: gap of unknown length
* 4568 6096: contig of 1529 bp in length
* 6097 7436: gap of unknown length
* 7437 7536: contig of 1240 bp in length ;
* 7537 8767: gap of unknown length
* 8768 8867: gap of unknown length
* 8868 10133: contig of 1266 bp in length
* 10134 10233: gap of unknown length
* 10234 11520: contig of 1287 bp in length
* 11521 11620: gap of unknown length
* 11621 13036: contig of 1416 bp in length
* 13037 13136: gap of unknown length
* 13137 14227: contig of 1091 bp in length
* 14228 14327: gap of unknown length
* 14328 15361: contig of 1034 bp in length
* 15362 15461: gap of unknown length
* 15462 16998: contig of 1537 bp in length
* 16999 17098: gap of unknown length
* 17099 18219: contig of 1121 bp in length
* 18220 18319: gap of unknown length

```







```
QY 241 AAGCGCTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCCGACCTGCTTTAT 300
|||||
Db 159506 AAGCCCTGGAGCTGTACCACCATCTACCTGATGAACCTGATGTGGCCGACCTGCTTTAC 159447
QY 301 GTGCT---ATTGCCCTTCCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGG 357
|||||
Db 159446 ATGCTGAGTCTGCGCTTCCTCATCATCACTACCTACCGGGGGTGACACCTTGGCTCTTTGGG 159387
QY 358 GAGCTGCTGCAAGCTGGTGCACTTCCCTGTTCTATATCAACCTTTACGGCAGCATCCTG 417
|||||
Db 159386 GAGACGCTGTGCAAGCTGGTGCGCTTCCCTGTTCTACCAACCTCTACAGCAGCATCCTG 159327
QY 418 CTGCTGACCTGCACTCTCTGTGCAACCACTTCTAGGTGTGTGCCACCACTGTGTGGCTG 477
|||||
Db 159326 CTGCTGACCTGCACTCTCTATATACCGCTTCTAGGAGTGTGCCACCACTGTGAGCTGCTG 159267
QY 478 CCTACCGGACCGCAGGAGCTGCTGGTGGGACACGACACCACTGGGGCCCTGGTGGTC 537
|||||
Db 159266 TCCCTCCGGACCGCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 159207
QY 538 CTCACGCTGCTGCCACACTGGCTTCTGCCACAGCACTACATCAATGCCAGATGATC 597
|||||
Db 159206 GTCCAATTGCTACCGACTCTGCTGCTTCTGCCACAGCACTATGTAATGGCCACGCTGATC 159147
QY 598 TGGTATGACATGACAGCCCAAGAGAAATTTTGATGGCTTTTTCCTACGGCATAGTTCTG 657
|||||
Db 159146 TGCTACGACACAGCCAGCCAGAGAAATTCGATCATTTCTTCGCCCTACAGCATGGTCTG 159087
QY 658 ACATTGCTGGCTT---TCTTTCCCTCTGGTGTGCTATTTTGGTGTGCTATTCACCTGATGGTC 715
|||||
Db 159086 ACATATCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 159027
QY 716 AGGAGCTGATCAAGCAGAGGAGAACCTCATGAGCAGGCAACACAGCCCGGAGCCAGG 775
|||||
Db 159026 AGGAGCTGACAAATCCAGCAGAGGAGCTCCCAACAGCAGCCAGCCCGGAGCCAGG 158967
QY 776 TCCATCCGGACCATCTCTACTGGTG 799
|||||
Db 158966 TCTATCCGGACCATCTCTGCTGGTG 158943
```

```
RESULT 15
AX244718 681 bp DNA linear PAT 28-SEP-2001
Sequence 47 from Patent WO0166750.
ACCESSION AX244718
VERSION AX244718.1 GI:15859597
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 681)
AUTHORS Vogeli, G. and Wood, L.S.
TITLE G protein-coupled receptors
JOURNAL Patent: WO 0166750-A 47 13-SEP-2001;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES Location/Qualifiers
source 1..681
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 163 a 171 c 171 g 176 t
ORIGIN
Query Match 36.6%; Score 396.4; DB 6; Length 681;
Best Local Similarity 99.7%; Pred. No. 1.8e-91;
Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 685 GGTCAATTTTGGTGTGCTATTACCTGATGTTGTCAGGAGCTGATCAAGCCAGGAGGACCT 744
|||||
Db 681 GGTCACTTTTGGTGTGCTATTACCTGATGTTGTCAGGAGCTGATCAAGCCAGGAGGACCT 622
```

```
QY 745 CATGAGGACAGGCAACACAGCCGAGCCAGGTCCATCCGACCATCTACTGTGTGTGG 804
|||||
Db 621 CATGAGGACAGGCAACACAGCCGAGCCAGGTCCATCCGACCATCTACTGTGTGTGG 562
QY 805 CCTCTTACACCTCTGTTTGTGCCCCCTTCATATCATCTGCTCCTTCTACCTCACCATCTG 864
|||||
Db 561 CCTCTTACACCTCTGTTTGTGCCCCCTTCATATCATCTGCTCCTTCTACCTCACCATCTG 502
QY 865 CTTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCCTACAGATATG 924
|||||
Db 501 CTTTCTGCTTCTCAGGACTGCCAGCTCTTGATGGCAGCCAGTGTGGCCTACAGATATG 442
QY 925 GAGGCCCTCTGGTGTGAGTGTGAGCAGCTGCCCTCAACCCAGTCTGTACTTTCTTTCAAGGG 984
|||||
Db 441 GAGGCCCTCTGGTGTGAGTGTGAGCAGCTGCCCTCAACCCAGTCTGTACTTTCTTTCAAGGG 382
QY 985 GGCAAAAATAGATCAGGCTCTCCAGAAACTGAGGCGAGAACAAAGTTGGGTGAGCATCCA 1044
|||||
Db 381 GGCAAAAATAGATCAGGCTCTCCAGAAACTGAGGCGAGAACAAAGTTGGGTGAGCATCCA 322
QY 1045 GCTGGGAGGAGAGATGCCCGAGGTTTGAACAGATCTGG 1082
|||||
Db 321 GCTGGGAGGAGAGATGCCCGAGGTTTGAACAGATCTGG 284
```

Search completed: May 26, 2003, 22:35:37  
Job time : 3961 secs





GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 20:12:01 ; Search time 311 Seconds  
(without alignments)  
7834.927 Million cell updates/sec

Title: US-10-023-586B-1  
Perfect score: 1082  
Sequence: 1 atgctgtccatttgcctcc.....ccagggttgacagatctgg 1082

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1082	100.0	3143	24 AAS17746	Human genomic clone
2	1051	97.1	1063	24 AAS98135	Human DNA for pote
3	1002.8	92.7	1076	24 AAD29667	Human G-protein co
4	996.4	92.1	1002	22 AAH73516	Human G-protein-co
5	850	78.6	850	24 AAH73516	Human p2y-like G p
6	510	47.1	510	22 AAH50975	Human nGPCR12 codi
7	396.4	36.6	681	22 AAS30774	Human cDNA encodin
8	370	34.2	585	22 AAS30783	Human cDNA encodin
9	227.8	21.1	2025	22 AAD04980	Human purinergic r

c	10	223.6	20.7	1163	22	AAD04981	Human purinergic r
	11	222	20.5	551	22	AAS42837	Human G Protein-Co
	12	202.2	18.7	1842	16	AAQ88134	Human P20 receptor
	13	202.2	18.7	2138	24	AAD32937	Mus musculus GPCR
	14	202.2	18.7	2138	24	AAD32944	Mus musculus GPCR
	15	178.6	16.5	1429	18	AAT74321	Human P2Y4 recepto
	16	170	15.7	234	22	AAS30782	Human cDNA encodin
	17	164.8	15.2	984	17	AAT18368	Human placenta G-p
	18	156.8	14.5	984	18	AAT70000	P2U purinergic rec
	19	156.8	14.5	984	18	AAT47730	Human placenta pur
	20	154	14.2	593	23	AAS66708	DNA encoding novel
	21	148.8	13.8	984	17	AAT18367	Mouse pancreas G-p
	22	105.8	9.8	6721	24	AAS18599	Purinergic recepto
	23	105.4	9.7	6721	24	AAS18600	Purinergic recepto
	24	97.8	9.0	1543	24	ABK12957	DNA sequence of mo
	25	97.4	9.0	1119	23	ABT98009	Non-endogenous hum
	26	95.6	8.8	1892	24	AAI72351	DEZ orphan recepto
	27	92	8.5	1477	16	AAQ84557	Murine C140 recept
	28	92	8.5	1477	17	AAT32036	Murine C140 recept
	29	92	8.5	2732	16	AAQ84559	Murine C140 recept
	30	92	8.5	2732	17	AAT32038	Murine C140 recept
	31	91.4	8.4	1176	23	ABT98024	Non-endogenous hum
	32	91.4	8.4	1634	14	AAQ45653	Human somatostatin
	33	90.6	8.4	9409	22	AAK76477	Human immune/haema
	34	89	8.2	1526	24	ABK94922	Human novel polynu
c	35	88.8	8.2	1001	24	AAS98131	Human DNA for pote
	36	88.4	8.2	450	21	AAA42673	Human secreted exp
	37	88.4	8.2	1020	21	AAK30601	Human G protein-co
	38	88.4	8.2	1020	21	AAK30718	DNA encoding human
	39	88.4	8.2	1020	24	ABK52848	Human cysteinyl le
	40	88.4	8.2	1901	15	AAQ66178	Seven transmembran
	41	88.4	8.2	1901	19	AAV18356	Human R12 seven tr
	42	88.4	8.2	1901	21	AAA91725	Human 7TM receptor
	43	88.4	8.2	1901	24	ABK54254	Human 7 transmembr
	44	88.4	8.2	2453	18	AAT44092	Human G-protein th
	45	87.4	8.1	1119	21	AAA30626	Human G protein-co

ALIGNMENTS

RESULT 1  
AAS17746  
ID AAS17746 standard; DNA; 3143 BP.

XX AC AAS17746;

XX DT 26-FEB-2002 (first entry)

XX Human genomic clone for P2Y-like G protein-coupled receptor.

XX Human; ds: P2Y-like G protein-coupled receptor; GPCR; COPD;  
XX Chronic obstructive pulmonary disease; nervous system disease;  
XX Parkinson's disease; multiple sclerosis; dementia; stroke;  
XX Alzheimer's disease; benign prostatic hyperplasia; urinary  
XX bacterial infection; fungal infection; protozoan infection;  
XX viral infection; pain; cancer; anorexia; bulimia; asthma;  
XX acute heart failure; hypotension; hypertension; osteoporosis; diabetes;  
XX angina pectoris; myocardial infarction; ulcer; inflammation; allergy;  
XX psychotic disorder; neurological disorder; anxiety; schizophrenia;  
XX manic depression; delirium; severe mental retardation; dyskinesia.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX CDS 520..2498  
XX /\*tag= a  
XX /product= "P2Y-like GPCR"

XX WO200185764-A2.

XX PD 15-NOV-2001.

XX XX

PF 09-MAY-2001; 2001WO-EP05244.  
XX  
PR 11-MAY-2000; 2000US-203582P.  
XX 21-FEB-2001; 2001US-269857P.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Ramakrishnan S;  
XX  
DR WPI: 2002-075242/10.  
XX P-PSDB; AAU11251.  
XX  
PT New polynucleotides for producing P2Y-like G protein-coupled receptors  
PT (GPCR) that are used for screening inhibitors or regulators of human  
PT P2Y-like GPCR, especially useful for treating pain, cancer or  
PT neurological disorders  
XX  
PS Disclosure: Fig 1: 114pp; English.  
XX  
CC The invention relates to an isolated polynucleotide encoding a P2Y-like  
CC G protein-coupled receptor (GPCR) polypeptide, its fragment,  
CC derivative or allele, a host cell containing an expression vector  
CC comprising the polynucleotide and screening for agents that regulate the  
CC GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR  
CC polypeptide, which may be employed for screening agents that inhibit or  
CC regulate human P2Y-like GPCR. The reagent or inhibitor of the human  
CC P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR  
CC disorders, particularly COPD (chronic obstructive pulmonary disease),  
CC peripheral or central nervous system disease (e.g. Parkinson's  
CC disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and  
CC many other diseases and disorders listed in the specification), benign  
CC prostatic hyperplasia or urinary incontinence. A pharmaceutical  
CC composition containing the modulators and/or regulators of P2Y-like  
CC GPCR is useful for modulating the activity of a P2Y-like GPCR.  
CC In particular, these are useful for treating, preventing or ameliorating  
CC infections (e.g. bacterial, fungal, protozoan or viral infections), pain,  
CC cancer, anorexia, bulimia, asthma, acute heart failure, hypotension,  
CC hypertension, osteoporosis, diabetes, angina pectoris, myocardial  
CC infarction, ulcers, inflammation, allergies, psychotic or neurological  
CC disorders (e.g. anxiety, schizophrenia, manic depression, delirium,  
CC severe mental retardation or dyskinesias). The present sequence is  
CC a genomic clone encoding the P2Y-like GPCR of the invention.  
XX  
SQ Sequence 3143 BP; 818 A; 771 C; 726 G; 828 T; 0 other;  
  
Query Match 100.0%; Score 1082; DB 24; Length 3143;  
Best Local Similarity 100.0%; Pred. No: 1.9e-284;  
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGCTGTCCATTTTGTCTTCCAGGGGAGCAGAGCGGGAGCGTGTGGAGCTCTG 60  
DB 436 ATGCTGTCCATTTTGTCTTCCAGGGGAGCAGAGCGGGAGCGTGTGGAGCTCTG 495  
  
61 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATACATCAGAGAACAA 120  
DB 496 CTCCTGGAGGAGCCTCCCGGACATGGAGAGGTGGACATGAATACATCAGAGAACAA 555  
  
121 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 180  
DB 556 GGTCTCTGCCAGTTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATC 615  
  
181 ATCTTTATCTAGGGTGGCCATTAATGGCACTGTCTTGTGGCACTCTCCGGCCCAACCC 240  
DB 616 ATCTTTATCTAGGGTGGCCATTAATGGCACTGTCTTGTGGCACTCTCCGGCCCAACCC 675  
  
241 AAGCGCTGAGCTGTGCCACCACTATCTGTGGAACCTGATGTGGCGGACCTGTTTAT 300  
DB 676 AAGCGCTGAGCTGTGCCACCACTATCTGTGGAACCTGATGTGGCGGACCTGTTTAT 735  
  
301 GTGCTATTGCCCTTCTCATCATCACTACTACTACATGATGACAGGTGGCCCTTCGGGGAG 360  
DB 736 GTGCTATTGCCCTTCTCATCATCACTACTACTACATGATGACAGGTGGCCCTTCGGGGAG 795

QY 361 CTGCTCTGCAAGCTGTGTGCACCTTCTCTTATATCAACCTTTACGCGCAGCATCTCTGCTG 420  
DB 796 CTGCTCTGCAAGCTGTGTGCACCTTCTCTTATATCAACCTTTACGCGCAGCATCTCTGCTG 855  
  
QY 421 CTGACCTGCATCTCTGTGTGCACCAAGTTCCTAGGTGTGTGCACCACTGTGTGCGTGCCTC 480  
DB 856 CTGACCTGCATCTCTGTGTGCACCAAGTTCCTAGGTGTGTGCACCACTGTGTGCGTGCCTC 915  
  
QY 481 TACCGGACCCGAGGAGCATGCTGGTGGCCACGACCACTGGGCGCTGTGTGCTC 540  
DB 916 TACCGGACCCGAGGAGCATGCTGGTGGCCACGACCACTGGGCGCTGTGTGCTC 975  
  
QY 541 CAGCTGTCTGCCACACTGGCCCTTCTCCACAGGACTACATCAATGCGCAGATGATCTGG 600  
DB 976 CAGCTGTCTGCCACACTGGCCCTTCTCCACAGGACTACATCAATGCGCAGATGATCTGG 1035  
  
QY 601 TATGACATGACCAAGAGAGAAATTTTGTGCGCTTTTGGCTACGGCATAGTCTTGACA 560  
DB 1036 TATGACATGACCAAGAGAGAAATTTTGTGCGCTTTTGGCTACGGCATAGTCTTGACA 1095  
  
QY 661 TTGTCTGGCTTTCTTCCCTCTTGGTCAATTTTGGTGTGCTATTCTACTCATGCTGTCAGGAG 720  
DB 1096 TTGTCTGGCTTTCTTCCCTCTTGGTCAATTTTGGTGTGCTATTCTACTCATGCTGTCAGGAG 1155  
  
QY 721 CCTGATCAAGCAGGAGAGAACTCATGAGGACAGCAACAGCCGAGCCAGGTCCTCAT 780  
DB 1156 CCTGATCAAGCAGGAGAGAACTCATGAGGACAGCAACAGCCGAGCCAGGTCCTCAT 1215  
  
QY 781 CCGGACCACTCTACTGCTGTGTGGCTCTTCCCTCTTGTGCTGCTTTGTGCGCTTCCATATCAC 840  
DB 1216 CCGGACCACTCTACTGCTGTGTGGCTCTTCCCTCTTGTGCTGCTTTGTGCGCTTCCATATCAC 1275  
  
QY 841 TCGCTCTCTTCTACCTCACCATCTGCTTTCTGCTTCTTCAGGACTGCCAGCTCTTGTATGGC 900  
DB 1276 TCGCTCTCTTCTACCTCACCATCTGCTTTCTGCTTCTTCAGGACTGCCAGCTCTTGTATGGC 1335  
  
QY 901 AGCCAGTGTGGCTTCAAGATATGAGGAGGCTCTGTGAGTGTGAGCAGTGTGCTCAACCC 960  
DB 1336 AGCCAGTGTGGCTTCAAGATATGAGGAGGCTCTGTGAGTGTGAGCAGTGTGCTCAACCC 1395  
  
QY 961 ACTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAAGTCTCTCCAGAACTGAGG 1020  
DB 1396 ACTCTGTACTTCTTCAAGGGGGGCAAAATAGAGTCAAGTCTCTCCAGAACTGAGG 1455  
  
QY 1021 CAGAACAAAGTGTGGTGTGAGCATCCAGCTGGGAGAGAGATGCCAGGGTTGAACAGATCT 1080  
DB 1456 CAGAACAAAGTGTGGTGTGAGCATCCAGCTGGGAGAGAGATGCCAGGGTTGAACAGATCT 1515  
  
QY 1081 GG 1082  
DB 1516 GG 1517  
  
RESULT 2  
AAS98135  
ID AAS98135 standard; DNA; 1063 BP.  
XX  
AC AAS98135;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Human DNA for potential G protein-coupled receptor #92.  
XX  
KW Human; G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;  
KW Alzheimer's disease; amyotrophic lateral sclerosis; asthma;  
KW atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;  
KW chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;  
KW depression; epilepsy; macular degeneration; lymphoma; melanoma;  
KW multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;  
KW psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;  
KW tuberculosis; cognition disorder; memory disorder; anorexia;  
KW hormonal release disorder; cardiovascular activity disorder;  
KW pain perception disorder; obesity; diabetes; obesity;

KW diabetes; hyperlipidaemia; stroke; gene therapy.

OS Homo sapiens.

PN WO200185791-A1.

XX 15-NOV-2001.

XX 11-MAY-2001; 2001WO-US15332.

XX 11-MAY-2000; 2000US-203217P.

PR 18-MAY-2000; 2000US-205945P.

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

XX Brown JP, Millier M, Burmer G, Fabre-Suver C, Pritchard D;

XX WPI; 2002-066595/09.

DR Novel G protein-coupled receptor polypeptides including galanin  
XX receptor polypeptides useful for identifying modulators that are useful  
PT for treating Alzheimer's disease, psoriasis, melanoma, multiple  
PT sclerosis, stroke

XX Claim 2; Page 98; 144pp; English.

XX The invention relates to an isolated polypeptide encoded by a  
CC nucleic acid molecule that is at least 80% identical to the G  
CC protein-coupled (GPCR) polynucleotides included in the specification.  
CC Also included are probes based on the GPCR sequences (including  
CC antisense sequences), a host cell comprising an expression vector comprising  
CC the GPCR sequence, antibodies raised against the polypeptides,  
CC and methods of identifying modulators of the polypeptides. The  
CC polypeptides are useful for identifying modulator compounds which  
CC function as modulators, activators, repressors, agonists or antagonists  
CC of the novel GPCR polypeptides including the GAL4 polypeptide. The  
CC antibodies and nucleic acid probes as described above can be used to  
CC detect the presence of the polypeptides and nucleic acids and are used to  
CC diagnose a variety of diseases or disorders in which GPCRs are involved  
CC e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma,  
CC atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy,  
CC chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease,  
CC depression, epilepsy, macular degeneration, lymphoma, melanoma,  
CC multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease,  
CC psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis,  
CC tuberculosis and many other diseases listed in the specification. The  
CC probes and antibodies are also useful for diagnosing cognition and memory  
CC disorders, anorexia, hormonal release disorders, cardiovascular activity  
CC disorders, pain perception disorders, obesity, diabetes, Alzheimer's  
CC disease. Preferably, compounds that decrease or increase  
CC the expression of galanin receptor (GAL4) can be used to treat obesity,  
CC diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is  
CC useful for treating the above mentioned disorders by gene therapy  
CC techniques. The present sequence is a novel GPCR polynucleotide of the  
CC invention.

XX SQ Sequence 1063 BP; 216 A; 315 C; 272 G; 260 T; 0 other;

Query Match 97.1%; Score: 1051; DB 24; Length 1063;  
Best Local Similarity 99.9%; Pred. No. 3.3e-276;  
Matches 1062; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 19 CCTTCCAGGGGAAGCAGAGCGGAGCGCTCGTGGAGCTCTGCTCTCGAGGAGGAGCGCTCC 78

Db 1 CCTTCCAGGGGAAGCAGAGCGGAGCGCTCGTGGAGCTCTGCTCTCGAGGAGGAGCGCTCC 60

QY 79 CGGACATGGAGAGGTGGACATATACATACAGGAACAAGTCTCTCTCCAGTTTCTCA 138

Db 61 CGGACATGGAGAGGTGGACATATACATACAGGAACAAGTCTCTCTCCAGTTTCTCA 120

QY 139 GAGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATCTAGGCGTG 198

Db 121 GAGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATCTAGGCGTG 180

QY 199 CCACTAAATGGCACTGTCTTTGTGGCACTCTCTGGGGCCAAACCAAGGCTGGAGCTGTGCC 258  
Db 181 CCACTAAATGGCACTGTCTTTGTGGCACTCTCTGGGGCCAAACCAAGGCTGGAGCTGTGCC 240  
QY 259 ACCACCTATCTGTGTGAACCTGATGGTGGCGACCTGCTTTATGTGTATTGCTTCCTTCCTC 318  
Db 241 ACCACCTATCTGTGTGAACCTGATGGTGGCGACCTGCTTTATGTGTATTGCTTCCTTCCTC 300  
QY 319 ATCATCACCTACTACTAGATGACAGTGGCCCTTCGGGGAGCTGCTCTCAAGCTGTGGT 378  
Db 301 ATCATCACCTACTACTAGATGACAGTGGCCCTTCGGGGAGCTGCTCTCAAGCTGTGGT 360  
QY 379 CACTTCTCTGTCTATATCAACCTTTACGGCAGCATCCTGCTGTGCTGACCTGCATCTCTGTG 438  
Db 361 CACTTCTCTGTCTATATCAACCTTTACGGCAGCATCCTGCTGTGCTGACCTGCATCTCTGTG 420  
QY 439 CACCACTTCTAGTGTGTGCCACCTGCTGTTCGCTGCGCTACCGGACCCCGAGGCGAT 498  
Db 421 CACCACTTCTAGTGTGTGCCACCTGCTGTTCGCTGCGCTACCGGACCCCGAGGCGAT 480  
QY 499 GCCTGGCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 558  
Db 481 GCCTGGCTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 540  
QY 559 GCCTTCTCCACACAGGCACTACATCAATGGCCAGATGATCTGGTATGACATGACAGCAGCAA 618  
Db 541 GCCTTCTCCACACAGGCACTACATCAATGGCCAGATGATCTGGTATGACATGACAGCAGCAA 600  
QY 619 GAGAAATTTTATCGGCTTTTGGCTACGGCATAGTTCTGTGATTTGTGCTTCTTCTTCTTTC 678  
Db 601 GAGAAATTTTATCGGCTTTTGGCTACGGCATAGTTCTGTGATTTGTGCTTCTTCTTCTTTC 660  
QY 679 CTCTTGGTTCATTTGGTGTGCTATTCACATGATGGTTCAGGAGCTGATCAAGCAGAGGA 738  
Db 661 CTCTTGGTTCATTTGGTGTGCTATTCACATGATGGTTCAGGAGCTGATCAAGCAGAGGA 720  
QY 739 GAACCTCATGAGGACAGGCAACACAGCCCGAGCAGGCTCCATCCGACCATCTTCTTCTTCTT 798  
Db 721 GAACCTCATGAGGACAGGCAACACAGCCCGAGCAGGCTCCATCCGACCATCTTCTTCTTCTT 780  
QY 799 GTGTGGCTCTTCACTCTGCTTTTGTGGCTTCCCATATACCTGCTGCTTCTTCTTCTTCTT 858  
Db 781 GTGTGGCTCTTCACTCTGCTTTTGTGGCTTCCCATATACCTGCTGCTTCTTCTTCTTCTT 840  
QY 859 CATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGGCAG-CCAGTGTGGCTTACA 917  
Db 841 CATCTGCTTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGGCAGCCAGCTGCTGCTTACA 900  
QY 918 AGATATGGAGGCTCTGCTGAGTGTGAGCAGCTGCCCTCAACCCAGTCTCTTCTTCTTCTT 977  
Db 901 AGATATGGAGGCTCTGCTGAGTGTGAGCAGCTGCCCTCAACCCAGTCTCTTCTTCTTCTT 960  
QY 978 CAAGGGGGGCAAAATAGATGATGAGTCTCCAGAACTAGGCGAGCAAGTGTGGTGA 1037  
Db 961 CAAGGGGGGCAAAATAGATGATGAGTCTCCAGAACTAGGCGAGCAAGTGTGGTGA 1020  
QY 1038 GCATCCAGCTGGGAGGAAGATGCCCGAGGTTGAACAGATCT 1080  
Db 1021 GCATCCAGCTGGGAGGAAGATGCCCGAGGTTGAACAGATCT 1063

RESULT 3

AAD29667

ID AAD29667 standard; cDNA; 1076 BP.

XX AAD29667;

XX AC

XX AC

DT 17-MAY-2002 (first entry)

XX Human G-protein coupled receptor (GCREC-1) cDNA.

XX Human; G-protein coupled receptor; GCREC-1; cell proliferative disorder;

KW neurologic; cardiovascular; gastrointestinal; autoimmune; inflammatory;  
KW metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;  
KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;  
KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;  
KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;  
KW transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;  
KW neurologic; neuroprotective; cardiant; immunosuppressive; anorectic;  
KW virulide; gene; ss.

OS Homo sapiens.

AA	Key	Location/Qualifiers
FH		57..1076
FT	CDS	

AA  
PN  
WO200210387-A2.

07-FEB-2002

25-JUL-2001: 2001WO-US23433.

27-JUL-2000: 2000US-221478P.

03-AUG-2000; 2000US-223268P.

21-AUG-2000; 2000US-227054P.

08-SEP-2000; 2000US-231121P.  
PR 13-SEP-2000; 2000US-232213P

[illegible]

PR 22-SEP-2000; 2000US-235146P.

DATE	DESCRIPTION	AMOUNT	CHECK NO.	BANK	INITIALS
1/1/20	DEPOSIT	100.00			
1/15/20	PAYROLL	50.00	101		
1/31/20	RENT	25.00	102		
2/1/20	DEPOSIT	150.00			
2/15/20	PAYROLL	50.00	103		
2/28/20	RENT	25.00	104		
3/1/20	DEPOSIT	200.00			
3/15/20	PAYROLL	50.00	105		
3/31/20	RENT	25.00	106		
4/1/20	DEPOSIT	250.00			
4/15/20	PAYROLL	50.00	107		
4/30/20	RENT	25.00	108		
5/1/20	DEPOSIT	300.00			
5/15/20	PAYROLL	50.00	109		
5/31/20	RENT	25.00	110		
6/1/20	DEPOSIT	350.00			
6/15/20	PAYROLL	50.00	111		
6/30/20	RENT	25.00	112		
7/1/20	DEPOSIT	400.00			
7/15/20	PAYROLL	50.00	113		
7/31/20	RENT	25.00	114		
8/1/20	DEPOSIT	450.00			
8/15/20	PAYROLL	50.00	115		
8/31/20	RENT	25.00	116		
9/1/20	DEPOSIT	500.00			
9/15/20	PAYROLL	50.00	117		
9/30/20	RENT	25.00	118		
10/1/20	DEPOSIT	550.00			
10/15/20	PAYROLL	50.00	119		
10/31/20	RENT	25.00	120		
11/1/20	DEPOSIT	600.00			
11/15/20	PAYROLL	50.00	121		
11/30/20	RENT	25.00	122		
12/1/20	DEPOSIT	650.00			
12/15/20	PAYROLL	50.00	123		
12/31/20	RENT	25.00	124		

PA (INCY-) INCYTE GENOMICS INC.  
yy

XX  
PI Thornton M. Patterson C. Lal P. Burford N. Yue H. Gandhi AR:

PI Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walia NK;

PI Hafalia AJA, Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L;

PI Graul RC, Warren BA, Lee EA, Ding L;

XX  
DR  
WPT: 2002-188744/2A

DR WFI, 2002-1007447.  
DR P-PSDB; AAE18640.

PT New human G-protein coupled receptor polypeptide for diagnosis,  
PT prevention and treatment of cell proliferative, neurological,  
PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and  
PT metabolic disorders -

PS Claim 5; Page 137-138; 150pp; English.

The invention relates to novel human G-protein coupled receptors (GCRC) and their encoding polynucleotides. GCRC is useful as an immunogen for preparing monoclonal and polyclonal antibodies. GCRC is useful for diagnosing, treating and preventing a cell proliferative disorder (e.g., hepatitis, psoriasis, cancer), a neurological disorder (e.g., epilepsy, Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial infarction), gastrointestinal disorder (e.g., gastritis, cirrhosis, Crohn's disease), an autoimmune/inflammatory disorder (e.g., acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma, rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity, osteoporosis), and viral infections. GCRC is useful in a number of drug screening techniques, and to analyse the proteome of a tissue or cell type. GCRC is useful for creating knockin humanised animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier or affected individuals, and as hybridization probes for mapping naturally occurring genomic sequences. GCRC is useful in Southern or northern analysis, dot blot or other membrane-based technologies, in PCR technologies, in dipstick, pin, multiformat enzyme linked immunosorbant (ELISA)-like assays, and in microarrays utilising fluids or tissues from patients to detect altered GCRC expression. The present sequence is human GCRC-1 cDNA.

Db 1017 CCAGTCTGTACTTCTTTCAAGGGGGCAAAATAGAGTACAGGCTCCTCCAGAAACTGA 1076

## RESULT 4

AAH73516  
ID AAH73516 standard; DNA; 1002 BP.

XX

XX

AC AAH73516;

XX

DT 25-SEP-2001 (first entry)

XX

XX Human G protein-coupled receptor GPRv71 coding sequence.

DE

XX Human; guanosine triphosphate binding protein-coupled receptor;

XX G protein-coupled receptor; GPRv8; GPRv12; GPRv16; GPRv21; GPRv40;

KW GPRv47; GPRv71; GPRv72; cancer; liver cirrhosis;

XX Alzheimer's disease; cytostatic; hepatotropic; nootropic;

XX neuroprotective; gene therapy; peptide therapy; ds.

XX

OS Homo sapiens.

XX

XX WO200148188-A1.

XX

XX 05-JUL-2001.

PD

XX 28-DEC-2000; 2000WO-JP09408.

XX

XX 28-DEC-1999; 99JP-0375152.

PR

XX 31-MAR-2000; 2000JP-010139.

XX

XX (HELI-) HELIX RES INST.

PA

XX Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;

PI Sugiyama T, Kishimoto T, Kanzaki K, Yasuda S, Inoue Y;

PI

XX WPI: 2001-425662/45.

DR

XX P-PSDB; AAG64125.

XX

XX New DNA encoding guanosine triphosphate binding protein coupled

PT receptors and their expression products for screening potential

PT anticancer and nootropic drugs and in diagnosis of these diseases

PT

XX Example 1; Page 142-143; 170pp; Japanese.

PS

XX The invention relates to nine human guanosine triphosphate binding

XX protein (G protein)-coupled receptors designated GPRv8, GPRv12, GPRv16,

CC GPRv21, GPRv40, GPRv47, GPRv71 and GPRv72, and to the

CC genes encoding them. These genes and proteins and antibodies against

CC the protein are useful in the treatment, prevention, diagnosis and

CC investigation of diseases associated with G protein-coupled receptors,

CC including cancer, cirrhosis of the liver and Alzheimer's disease.

CC The present sequence is a G protein-coupled receptor coding sequence

CC of the invention.

XX

XX Sequence 1002 BP; 207 A; 293 C; 249 G; 253 T; 0 other;

SQ

Query Match 92.1%; Score 996.4; DB 22; Length 1002;

Best Local Similarity 99.9%; Pred. No. 2.5e-261;

Matches 997; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 85 ATGGAGAAGGTGCATGATATACATACAGGACACAGGCTCTGCCAGTTCTCAGAGAAG 144

DB 1 ATGGAGAAGGTGCATGATATACATACAGGACACAGGCTCTGCCAGTTCTCAGAGAAG 60

QY 145 TACAAGCAAGTCTACCTCTCCCGGCTACAGTATCATCTTTATCTAGGCTGCCACTA 204

DB 61 TACAAGCAAGTCTACCTCTCCCGGCTACAGTATCATCTTTATCTAGGCTGCCACTA 120

QY 205 AATGGCACTGTCTTTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACCACC 264

DB 121 AATGGCACTGTCTTTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCTGTGCCACCACC 180

QY 265 TATCTGGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCTATGTGCCCTTCCTCATCATC 324

Db 181 TATCTGGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCTATTGCGCTTCCTCATCATC 240

QY 325 ACCTACTCAGTACAGAGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACCTTC 384

Db 241 ACCTACTCAGTACAGAGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACCTTC 300

QY 385 CTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGCACTCTCTGTGCACAG 444

Db 301 CTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGCACTCTCTGTGCACAG 360

QY 445 TTCTTAGTGTGTGCCACCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504

Db 361 TTCTTAGTGTGTGCCACCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

QY 505 CTGGGACACAGCAGCACCCTGGGCTGGTGGTCTCCAGTGTGCTGCCACACACACCTTC 564

Db 421 CTGGGACACAGCAGCACCCTGGGCTGGTGGTCTCCAGTGTGCTGCCACACACACCTTC 480

QY 565 TCCACACAGGACTACATCAATGCGCAGATGATCTGGTATGACATGACACGACCCCAAGAGAAT 624

Db 481 TCCACACAGGACTACATCAATGCGCAGATGATCTGGTATGACATGACACGACCCCAAGAGAAT 540

QY 625 TTTGATCGGCTTTTTCCTACGGCATAGTCTGACATTTGCTGCTGCTTCTTCTCCCTCCTT 684

Db 541 TTTGATCGGCTTTTTCCTACGGCATAGTCTGACATTTGCTGCTGCTTCTTCTCCCTCCTT 600

QY 685 GGTCAATTTTGGTGTCTATCTGATGTCAGGACCTGATCAAGCAGCAGAGAACCT 744

Db 601 GGTCAATTTTGGTGTCTATCTGATGTCAGGACCTGATCAAGCAGCAGAGAACCT 660

QY 745 CATGAGCAGGCAACACAGCCGAGCCAGCTCCATCCGAGCAGCTCTACTGTGTGTGG 804

Db 661 CATGAGCAGGCAACACAGCCGAGCCAGCTCCATCCGAGCAGCTCTACTGTGTGTGG 720

QY 805 CCTCTTCACTCTGTTTGTGGCTTCCATATCATCTGCTGCTCTTCTACCTACCATCTG 864

Db 721 CCTCTTCACTCTGTTTGTGGCTTCCATATCATCTGCTGCTCTTCTACCTACCATCTG 780

QY 865 CTCTTCTGCTTCTCAGGACTGCCAGCTCTTGTGAGCAGCAGCTGCTCTACTGATATG 924

Db 781 CTCTTCTGCTTCTCAGGACTGCCAGCTCTTGTGAGCAGCAGCTGCTCTACTGATATG 840

QY 925 GAGGCTCTGCTGAGTGTGAGCAGCTGCCATCAACCCAGCTCTCTACTTTCTTCAAGGG 984

Db 841 GAGGCTCTGCTGAGTGTGAGCAGCTGCCATCAACCCAGCTCTCTACTTTCTTCAAGGG 900

QY 985 GGCAGAAATAGAGTCAAGCTCTCCAGAACTCAGGAGCAGAGAACAAAGTTGGTGAGCATCCA 1044

Db 901 GGCAGAAATAGAGTCAAGCTCTCCAGAACTCAGGAGCAGAGAACAAAGTTGGTGAGCATCCA 960

QY 1045 GCTGGGAGGAAGAGATGCCAGGGTTGAACAGATCTGG 1082

Db 961 GCTGGGAGGAAGAGATGCCAGGGTTGAACAGATCTGG 998

RESULT 5

AAH17747

ID AAH17747 standard; DNA; 850 BP.

XX

AC AAH17747;

XX

DT 26-FEB-2002 (first entry)

XX

XX Human P2Y-like G protein-coupled receptor exon 1.

DE

XX Human; ds; P2Y-like G protein-coupled receptor; GPCR; COPD;

KW chronic obstructive pulmonary disease; nervous system disease;

KW Parkinson's disease; multiple sclerosis; dementia; stroke;

KW Alzheimer's disease; benign prostatic hyperplasia; urinary incontinence;

KW bacterial infection; fungal infection; protozoan infection;

KW viral infection; pain; cancer; anorexia; bulimia; asthma; exon 1;

KW acute heart failure; hypotension; hypertension; osteoporosis; diabetes;

angina pectoris; myocardial infarction; ulcer; inflammation; allergy;  
psychotic disorder; neurological disorder; anxiety; schizophrenia;  
manic depression; delirium; severe mental retardation; dyskinesia.  
Homo sapiens.  
WO200185764-A2.  
15-NOV-2001.  
09-MAY-2001; 2001WO-EP05244.  
11-MAY-2000; 2000US-203582P.  
21-FEB-2001; 2001US-269857P.  
(FARB ) BAYER AG.  
Ramakrishnan S;  
WPI; 2002-075242/10.  
New polynucleotides for producing P2Y-like G protein-coupled receptors  
(GPCR) that are used for screening inhibitors or regulators of human  
P2Y-like GPCR, especially useful for treating pain, cancer or  
neurological disorders  
Claim 1; Fig 4; 114pp; English.  
The invention relates to an isolated polynucleotide encoding a P2Y-like  
G protein-coupled receptor (GPCR) polypeptide, its fragment,  
derivative or allele, a host cell containing an expression vector  
comprising the polynucleotide and screening for agents that regulate the  
GPCR activity. The polynucleotide is useful for producing P2Y-like GPCR  
polypeptide, which may be employed for screening agents that inhibit or  
regulate human P2Y-like GPCR. The reagent or inhibitor of the human  
P2Y-like GPCR is useful for treating or ameliorating P2Y-like GPCR  
disorders, particularly COPD (chronic obstructive pulmonary disease),  
peripheral or central nervous system disease (e.g. Parkinson's  
disease, multiple sclerosis, dementia, stroke, Alzheimer's disease and  
many other diseases and disorders listed in the specification), benign  
prostatic hyperplasia or urinary incontinence. A pharmaceutical  
composition containing the modulators and/or regulators of P2Y-like  
GPCR is useful for modulating the activity of a P2Y-like GPCR.  
In particular, these are useful for treating, preventing or ameliorating  
infections (e.g. bacterial, fungal, protozoan or viral infections), pain,  
cancer, anorexia, bulimia, asthma, acute heart failure, hypotension,  
hypertension, osteoporosis, diabetes, angina pectoris, myocardial  
infarction, ulcers, inflammation, allergies, psychotic or neurological  
disorders (e.g. anxiety, schizophrenia, manic depression, delirium,  
severe mental retardation or dyskinesias). The present sequence is  
exon 1 of the P2Y-like GPCR of the invention.  
Sequence 850 BP; 165 A; 262 C; 201 G; 222 T; 0 other;  
Query Match 78.68; Score 850; DB 24; Length 850;  
Best Local Similarity 100.08; Pred. No. 1.8e-221;  
Matches 850; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 85 ATGGAGAAGGTGGACATCAATACATCACAGGAACAGGTCTCTGCAGTCTTCACAGAGAAG 144  
DB 1 ATGGAGAAGGTGGACATCAATACATCACAGGAACAGGTCTCTGCAGTCTTCACAGAGAAG 60  
QY 145 TACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATCTAGGGCTGCCACTA 204  
DB 61 TACAAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATCTAGGGCTGCCACTA 120  
QY 205 ATGGCACTGTCTTGTGGCACTCTCTGGGGCCCAACCAAGCGCTGGAGCTGGCCACCACC 264  
DB 121 AATGGCACTGTCTTGTGGCACTCTCTGGGGCCCAACCAAGCGCTGGAGCTGGCCACCACC 180  
QY 265 TATCTGGTGAACCTCATGGTGGCCGACCTGCTTTATGTGCTATTGCTCCCTTCTCATCATC 324  
DB 181 TATCTGGTGAACCTCATGGTGGCCGACCTGCTTTATGTGCTATTGCTCCCTTCTCATCATC 240

QY 325 ACCTACTCAGTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACCTTC 384  
DB 241 ACCTACTCAGTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACCTTC 300  
QY 385 CTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTGCTGACCTGATCTCTGTGTGCACCA 444  
DB 301 CTGTTCTATATCAACCTTTACGGCAGCATCTCTGCTGCTGACCTGATCTCTGTGTGCACCA 360  
QY 445 TTCTTAGTGTGTGCCACCCACTGTGTTGCTGCTCCCTACCGACCCGAGGAGCATGCTGCTGG 504  
DB 361 TTCTTAGTGTGTGCCACCCACTGTGTTGCTGCTCCCTACCGACCCGAGGAGCATGCTGCTGG 420  
QY 505 CTGGCACACAGCACACCTGGGCGCTGTGTGCTCTCCAGCTGCTGCCACACTGSCCTTTC 564  
DB 421 CTGGCACACAGCACACCTGGGCGCTGTGTGCTCTCCAGCTGCTGCCACACTGSCCTTTC 480  
QY 565 TCCACACGGGACTACATCAATGGCCAGATGATCTGGTATGACATGACATGACCAAGAGAA 624  
DB 481 TCCACACGGGACTACATCAATGGCCAGATGATCTGGTATGACATGACCAAGAGAA 540  
QY 625 TTTGATCGGCTTTTTCCTACGGCAGCATGTTCTGACATTTGCTGGCTTTCTTTCCCTCCTT 684  
DB 541 TTTGATCGGCTTTTTCCTACGGCAGCATGTTCTGACATTTGCTGGCTTTCTTTCCCTCCTT 600  
QY 685 GGTCAATTTTGGTGTGCTATTCACATGATGGTCAGGAGCCTGATCAAGCCAGAGGAGAACT 744  
DB 601 GGTCAATTTTGGTGTGCTATTCACATGATGGTCAGGAGCCTGATCAAGCCAGAGGAGAACT 660  
QY 745 CATGAGGACAGCAACACAGCCCGAGCCAGGTCCATCCGGACCATCCTACTGTTGTGTGG 804  
DB 661 CATGAGGACAGCAACACAGCCCGAGCCAGGTCCATCCGGACCATCCTACTGTTGTGTGG 720  
QY 805 CCTCTTACCCCTCTGTTTGTGCTTCCATATCACTCGCTCTCTTCTACCTCACCACATCTG 864  
DB 721 CCTCTTACCCCTCTGTTTGTGCTTCCATATCACTCGCTCTCTTCTACCTCACCACATCTG 780  
QY 865 CTTTCTGCTTTTCTCAGGACTGCCAGCTCTTGTATGACGCCAGTGTGGCTCAAGAGATATG 924  
DB 781 CTTTCTGCTTTTCTCAGGACTGCCAGCTCTTGTATGACGCCAGTGTGGCTCAAGAGATATG 840  
QY 925 GAGGCTCTCTG 934  
DB 841 GAGGCTCTCTG 850  
RESULT 6  
AAH50975  
ID AAH50975 standard; DNA; 510 BP.  
XX AC AAH50975;  
XX DT 28-AUG-2001 (first entry)  
XX DE Human nPCR12 coding sequence.  
XX KW G protein-coupled receptor; nPCR; seven transmembrane receptor;  
XX KW signal transduction; schizophrenia; thyroid disorder; renal failure;  
XX KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;  
XX KW cardiovascular disease; proliferative disorder; hormonal disorder;  
XX KW neurological disorder; neuronal disorder; Alzheimer's disease; cancer;  
XX KW attention deficit-hyperactivity disorder/attention deficit disorder;  
XX KW Parkinson's disease; migraine; senile dementia; inflammatory disease;  
XX KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
XX KW neuroprotective; ds.  
XX OS Homo sapiens.  
XX PN WO200136473-A2.  
XX PD 25-MAY-2001.  
XX PF 16-NOV-2000; 2000WO-US31581.

XX 16-NOV-1999; 99US-0165838.  
PR 17-NOV-1999; 99US-0166071.  
PR 19-NOV-1999; 99US-0166678.  
PR 28-DEC-1999; 99US-0173396.  
PR 22-FEB-2000; 2000US-0184129.  
PR 28-FEB-2000; 2000US-0185421.  
PR 28-FEB-2000; 2000US-0185554.  
PR 02-MAR-2000; 2000US-0186530.  
PR 03-MAR-2000; 2000US-0186811.  
PR 07-MAR-2000; 2000US-0188114.  
PR 19-MAR-2000; 2000US-0190310.  
PR 21-MAR-2000; 2000US-0190800.  
PR 20-APR-2000; 2000US-0198568.  
PR 02-MAY-2000; 2000US-0201190.  
PR 08-MAY-2000; 2000US-0203111.  
PR 25-MAY-2000; 2000US-0207094.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
PA Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;  
PI Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;  
XX WPI; 2001-389826/41.  
DR P-PSDB; AAG80935.  
XX New G protein-coupled receptor (ngPCR-x) and its encoding  
PT polynucleotide useful for diagnosing and treating e.g. schizophrenia -  
PS Claim 4; Page 78; 261pp; English.  
XX The present invention relates to novel G protein-coupled receptors  
CC (ngPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,  
CC 28, 31-38, 40, 41, 53-60) and their coding sequences. The present  
CC sequence is the coding sequence for one such G protein-coupled receptor.  
CC GPCR are also known as seven transmembrane receptors and function in  
CC signal transduction. The ngPCRx coding sequences are useful for  
CC screening a human to diagnose a disorder affecting the brain or a genetic  
CC predisposition, specifically schizophrenia. ngPCRx are useful for  
CC identifying compounds useful for treating schizophrenia. Detection of  
CC ngPCRx in a sample is useful as a diagnostic tool for diseases or  
CC disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,  
CC CNS disorders, infections such as HIV-1, metabolic and cardiovascular  
CC diseases, proliferative disorders and hormonal disorders. Modulators of  
CC ngPCRx activity have the utility for treating neurological disorders,  
CC including schizophrenia, ADHD/ADD (attention deficit-hyperactivity  
CC disorder/attention deficit disorder), and neuronal disorders such as  
CC Alzheimer's disease, Parkinson's disease, migraine and senile dementia.  
CC Additional disorders include inflammatory conditions (e.g. Crohn's  
CC disease), rheumatoid arthritis, autoimmune disorders, cancers,  
CC respiratory ailments such as asthma, and inflammatory diseases e.g.  
CC inflammatory bowel disease.  
XX Sequence 510 BP; 92 A; 156 C; 126 G; 136 T; 0 other;  
SQ Query Match 47.1%; Score 510; DB 22; Length 510;  
Best Local Similarity 100.0%; Pred. No. 6.6e-129;  
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 247 TGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCCGACCTGTTTATGTGCTA 306  
DB 1 TGGAGCTGTGCCACCACTATCTGGTGAACCTGATGGTGGCCGACCTGTTTATGTGCTA 60,  
QY 307 TTGGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 366  
DB 61 TTGGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTC 120,  
QY 367 TGCAAGCTGGTGCACTTCTCTTATATCAACCTTTACGGCAGCATCTGCTGCTGACC 426  
DB 121 TGCAAGCTGGTGCACTTCTCTTATATCAACCTTTACGGCAGCATCTGCTGCTGACC 180  
QY 427 TGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTTTCGCTGCCCTACCGG 486  
DB 181 TGCATCTCTGTGCACCACTTCTAGGTGTGTGCCACCACTGTTTCGCTGCCCTACCGG 240  
QY 487 ACCCGCAGGATGCTGGCTGGGCACACGACACCACTGGGCCCTGGTGGTCTCCAGCTG 546  
DB 241 ACCCGCAGGATGCTGGCTGGGCACACGACACCACTGGGCCCTGGTGGTCTCCAGCTG 300  
QY 547 CTGCCACACTGGCCCTTCTCCACACGGGACTACATCAATGGCCAGATGATCTGGTATGAC 606  
DB 301 CTGCCACACTGGCCCTTCTCCACACGGGACTACATCAATGGCCAGATGATCTGGTATGAC 360  
QY 607 ATGACGAGCCAGAGAAATTTTGATCGGCTTTTGGCCCTACGGCATAGTTCTGACATGTCT 666  
DB 361 ATGACGAGCCAGAGAAATTTTGATCGGCTTTTGGCCCTACGGCATAGTTCTGACATGTCT 420  
QY 667 GGCTTTCTTCCCTCTCTGGTCAATTTTGGTGTGCTATTCTACTGATGTCAGGAGCCTGAT 726  
DB 421 GGCTTTCTTCCCTCTCTGGTCAATTTTGGTGTGCTATTCTACTGATGTCAGGAGCCTGAT 480  
QY 727 CAAGCCAGAGGAGAACCTCATGAGGACAGG 756  
DB 481 CAAGCCAGAGGAGAACCTCATGAGGACAGG 510  
XX AAS30774 standard; cDNA; 681 BP.  
XX AAS30774/C  
AC AAS30774;  
XX 04-DEC-2001 (first entry)  
DE Human cDNA encoding G protein-coupled receptor ngPCR-75.  
XX Human: G protein-coupled receptor; ngPCR-x; as; antiviral; analgesic;  
KW cytosolic; cardiac; antidiabetic; anorectic; hypotensive; hypertensive;  
KW antiparkinsonian; nootropic; neuroprotective; antidepressant;  
KW viral infection; HIV-1; human immunodeficiency virus; HIV-2; pain;  
KW cancer; metabolic disease; cardiovascular disease; type 2 diabetes;  
KW obesity; anorexia; hypotension; hypertension; myocardial infarction;  
KW atherosclerosis; Parkinson's disease; psychosis; neurological disorder;  
KW schizophrenia; migraine; major depression; anxiety; mental disorder;  
KW manic depression; dyskinesia; Huntington's disease; Tourette's Syndrome.  
OS Homo sapiens.  
XX WO200166750-A2.  
XX 13-SEP-2001.  
XX 08-MAR-2001; 2001WO-US07322.  
PR 08-MAR-2000; 2000US-0187581.  
PR 08-MAR-2000; 2000US-0187582.  
PR 08-MAR-2000; 2000US-0187714.  
PR 08-MAR-2000; 2000US-0187715.  
PR 08-MAR-2000; 2000US-0187825.  
PR 08-MAR-2000; 2000US-0187828.  
PR 08-MAR-2000; 2000US-0187829.  
PR 08-MAR-2000; 2000US-0187830.  
PR 08-MAR-2000; 2000US-0187833.  
PR 08-MAR-2000; 2000US-0187874.  
PR 08-MAR-2000; 2000US-0187930.  
PR 08-MAR-2000; 2000US-0188049.  
PR 08-MAR-2000; 2000US-0189294.  
PR 08-MAR-2000; 2000US-0189299.  
PR 08-MAR-2000; 2000US-0187928.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
PA Vogeli G, Wood LS;  
XX WPI; 2001-536778/59.  
DR P-PSDB; AAU19205.





Db 61 TAACCTCATGAGGTACAGGCAACACAGCCGAGCCAGGTCCATCCGGGACCATCCTACTG 120  
QY 797 GTGTGTCGCTTTCACCCCTCTGTTGTGTCCTTCCATATCATCTCCTCTTACTCTC 856  
Db 121 GTGTGTCGCTTTCACCCCTCTGTTGTGTCCTTCCATATCATCTCCTCTTACTCTC 180  
QY 857 ACCATCTGCTTTCTGCTTCTCAGGACTGCGAGCTCTTGATGGCAGCCAGTGTGGCCTAC 916  
Db 181 ACCATCTGCTTTCTGCTTCTCAGGACTGCGAGCTCTTGATGGCAGCCAGTGTGGCCTAC 240  
QY 917 AAGATATGGAGCCCTCTGCTGAGTGTGAGCAGCTGCTCAACCAGTCTCTGCTACTTTCTT 976  
Db 241 AAGATATGGAGCCCTCTGCTGAGTGTGAGCAGCTGCTCAACCAGTCTCTGCTACTTTCTT 300  
QY 977 TCAAGGGGGGCAAAATAGAGTCAAGGCTCTCTCCAGAAACTGAGCGACAACAAGTTGGTG 1036  
Db 301 TCAAGGGGGGCAAAATAGAGTCAAGGCTCTCTCCAGAAACTGAGCGACAACAAGTTGGTG 360  
QY 1037 AGCATCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCTGG 1082  
Db 361 AGCATCAGCTGGGAGGAGAGATGCCAGGGTTGAACAGATCTGG 406

RESULT 9

AAD04980  
ID AAD04980 standard; DNA: 2025 BP.  
XX  
AC AAD04980;  
XX  
DT 17-JUL-2001 (first entry)  
XX  
DE Human purinergic receptor P2Y2 DNA.  
XX  
KW Human; purinergic receptor; P2Y2; antidepressant; vulnerary; hypotensive;  
KW anticonvulsant; antiinflammatory; therapy; nervous system disorder;  
KW epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;  
KW irritable bowel disorder; reproductive system disorder; hypertension;  
KW peripheral vascular disease; immune system disorder; chronic bronchitis;  
KW premature ejaculation; asthma; neuromuscular disease; ds.  
XX  
OS Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX CDS 246..1379  
XX FT /\*tag= a  
XX FT /product= "Human P2Y2 receptor protein"  
XX FT complement (246..263)  
XX FT primer\_bind  
XX FT /\*tag= b  
XX FT /bound\_moiety= "Sense primer"  
XX FT 1362..1379  
XX FT /\*tag= c  
XX FT /bound\_moiety= "Antisense primer"  
XX  
XX US6214581-B1.  
XX  
XX 10-APR-2001.  
XX  
XX 13-NOV-1998; 98US-0191136.  
XX  
XX 16-JAN-1998; 98US-0071298.  
XX 16-JAN-1998; 98US-0071669.  
XX 16-JAN-1998; 98US-0008185.  
XX 16-JAN-1998; 98US-0008526.  
XX  
XX (ABBO ) ABBOTT LAB.  
XX  
XX Lynch KJ, Burgard EC, Van Biesen T;  
XX  
XX WPI: 2001-315459/33.  
XX P-PSDB: AAE01143.  
XX  
XX Novel isolated polynucleotide encoding human purinergic P2X3 receptor

PT polypeptide useful for identifying potentially therapeutic compounds  
PT that modulate or otherwise interact with P2X containing receptors -  
XX  
XX Example 14; Fig 11A; 53pp; English.  
XX  
CC The present sequence is a DNA encoding human wild type P2X4 receptor  
CC protein. P2X receptors are ligand-gated ion channels while P2Y receptors  
CC operate generally through a G-protein coupled system. P2X purinoreceptor  
CC drugs are potential therapeutic agents in several disorders including  
CC central nervous system or peripheral nervous system conditions, e.g.; the  
CC epilepsy, pain, depression, neurodegenerative disorders, disorders of the  
CC skeletal muscle such as neuromuscular diseases, disorders of the  
CC reproductive system, asthma, peripheral vascular disease, hypertension,  
CC immune system disorders, irritable bowel disorder, premature ejaculation,  
CC cystic fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the  
CC activity of extracellular nucleotide triphosphates to regulate chloride  
CC secretion in human airway epithelia.  
XX  
SQ Sequence 2025 BP; 388 A; 617 C; 595 G; 425 T; 0 other;  
Query Match 21.1%; Score 227.8; DB 22; Length 2025;  
Best Local Similarity 56.4%; Pred No. 9.9e-52;  
Matches 486; Conservative 0; Mismatches 367; Indels 8; Gaps 3;  
QY 127 TGCAGTTCTCAGAAAGTACAAAGAGTCTACCTCTCCTGGCTACAGTATCATCTTT 186  
Db 318 TGCCTTCAACGAGGAGCTTCAAGTACGTCTGCTGCTGTCTCTACGGCGTGTGTGC 377  
QY 187 ATCCTAGGGTGCACCTAAATGGCAGTGTCTTTGTGGCACTCTCTGGGGCCAAACGAGGC 246  
Db 378 GTGCTTGGGCTGTGCTGAACGGCGTGGCGCTCTACATCTTCTTGTGCGCGCTCAAGACC 437  
QY 247 TGGAGCTGTGCCACCACTATCTGTGAACCTGATGTGGCGGACCTGCTTTATGTGCTA 306  
Db 438 TGGAACTGCTCCACCATATATATGTTCCACTGGCTGTGTCTGATGACTGTTCGGGCC 497  
QY 307 T---TGGCCTTCTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGAGGTG 363  
Db 498 TCCCTGCGGCTGCTGGTCTATTACTACGCGCGGCGGACCACTGGCCCTTCAGCACGGTG 557  
QY 364 CTCTGCAAGCTGTGTCACCTTCTCTTATATCAACCTTTACGGCAGCATCTCTGCTGCT 423  
Db 558 CTCTGCAAGCTGTGTCGCTTCTCTTCTACACCAACCTTTACTGCAGCATCTCTCTCTC 617  
QY 424 ACCTGCAATCTCTGTGACCACTTCTTAGTGTGTGCCACCACTGTGTTCCTGCTCCCTAC 483  
Db 618 ACCTGCAATGAGCTGTGACCGGTGTCTGGGGCTCTTAGACCTCTGCGCTCCTCGGCTG 677  
QY 484 CGGACCCGCGAGGATGCTGCTGGGCAACCACTACCTGGGCGCTTGGTGGTCTCTCCAG 543  
Db 678 GGCGGGCGCGCTACGCTGCGCGGTGGCGGGCGGTGTGGGTGTGGTGTGGCTGCTGCTG 737  
QY 544 CTGCTGCCCACATGCGCTTCTCCACACGAGCTACATCAATGCCAGCATGATCTGTTAT 603  
Db 738 CAGCCCCCGGTGCTCTACTTTGTACCAACCACTGCGGGGCGCGCTGAACCTGCCAC 797  
QY 604 GACATGACAGCCCAAGAGAAATTTGATCGGCTTTTGTCTACGCACTAGTTCTGACATTG 663  
Db 798 GACACCTGGGACCCGAGCTCTTTCAGCGCTTCTGGCTGCTACAGCTCAGTCTGCTGGC 857  
QY 664 TCTGGCTT--TCTTTCCCTCTCTTGGTCTATTTGGTGTGCTATTCACATGATGGTCAGGAGC 721  
Db 858 CTGCTCTTCGCGGTGCCCTTTTGGCGTCTATCTCTCTGTTACGTGCTCATGCTCGCGCA 917  
QY 722 CTGATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACAGCCGCGAGCCAGGTCTCATC 781  
Db 918 CTGCTAAAGCCAGCCTACGGGACCTTCGGGGCGCTCTCCCTAGGGCAAGCGCAAGTCCGCTG 977  
QY 782 CGGACCATCTACTGCTGTGGCTCTTCCACCTCTGTTTGTGGCTTCCCATATCACT 841  
Db 978 CGCACCATGCGCGGTGGTGTGCTGCTTCCGCTCTGCTTCTGCTTCCCATCTCCAGTCACC 1037  
QY 842 CGCTCCTTCTACCTCACCATCTGCTTCTGCTTTCTCAGGACTGCCAGCTCTTGTATGCA 901

Db 1038 CGCACCCTCTACTACTCTCCCTCCGCT---CGCTGGACCTCAGCTGCCACACCTCAACGCC 1094  
 Qy 902 GCAGGTGGCTTACAGATATGGAGGCTCTGGTGAGTGTGAGCAGCTGCCTCAACCCA 961  
 Db 1095 ATCAACATGGCTTACAGGTATCCCGGCGCTGGCGCAGTGTCAACAGTTGCCTTGACCCC 1154  
 Qy 962 GTCCTGACTTCTTCTTCAAGG 982  
 Db 1155 GTGCTCTACTCTCTGGCTGGG 1175

## RESULT 10

AAD04981

ID AAD04981 standard; cDNA; 1163 BP.

AC AAD04981;

XX 17-JUL-2001 (first entry)

XX Human purinergic receptor P2Y2 cDNA.

XX Human: purinergic receptor; P2Y2; antidepressant; vulnerary; hypotensive;  
 anticonvulsant; antiinflammatory; therapy; nervous system disorder;  
 epilepsy; pain; depression; neurodegenerative disorder; cystic fibrosis;  
 irritable bowel disorder; reproductive system disorder; hypertension;  
 peripheral vascular disease; immune system disorder; chronic bronchitis;  
 premature ejaculation; asthma; neuromuscular disease; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FH primer\_bind complement (1..32)

FT /\*tag= a

FT /\*note= "Sense primer"

FT CDS 15..1148

FT /\*tag= b

FT /product= "protein encoded by human P2Y2 cDNA"

FT /transl\_except= (pos:786..788, aa:Phe)

FT misc\_feature 787

FT /\*tag= c

FT /\*note= "Represented in the specification as F"

FT primer\_bind 1131..1163

FT /\*tag= d

FT /bound\_moiety= "Antisense primer"

XX US6214581-B1.

XX 10-APR-2001.

XX 13-NOV-1998; 98US-0191136.

XX 16-JAN-1998; 98US-0071298.

XX 16-JAN-1998; 98US-0071669.

XX 16-JAN-1998; 98US-0008185.

XX 16-JAN-1998; 98US-0008526.

XX (ABBO ) ABBOTT LAB.

XX Lynch KJ, Burgard EC, Van Biesen T;

XX WPI: 2001-315459/33.

XX P-PSDB; AAE01114.

XX Novel isolated polynucleotide encoding human purinergic P2X3 receptor  
 polypeptide useful for identifying potentially therapeutic compounds  
 that modulate or otherwise interact with P2X containing receptors -

XX Example 14; Fig 12A; 53pp; English.

XX The present sequence is a cDNA encoding human P2Y2 receptor protein. P2X  
 CC receptors are ligand-gated ion channels while P2Y receptors operate  
 CC generally through a G-protein coupled system. P2X purinoreceptor drugs

CC are potential therapeutic agents in several disorders including central  
 CC nervous system or peripheral nervous system conditions, e.g., epilepsy,  
 CC pain, depression, neurodegenerative disorders, disorders of the skeletal  
 CC muscle such as neuromuscular diseases, disorders of the reproductive  
 CC system, asthma, peripheral vascular disease, hypertension, immune system  
 CC disorders, irritable bowel disorder, premature ejaculation, cystic  
 CC fibrosis and chronic bronchitis. P2Y purinoreceptors mediate the activity  
 CC of extracellular nucleotide triphosphates to regulate chloride secretion  
 CC in human airway epithelia.

XX Sequence 1163 BP; 177 A; 403 C; 332 G; 250 T; 1 other;

Query Match 20.7%; Score 223.6; DB 22; Length 1163;

Best Local Similarity 56.1%; Pred. No. 1.le-50;

Matches 483; Conservative 0; Mismatches 370; Indels 8; Gaps 3;

Qy 127 TGGCAGTTCTCAGAGCAAGTACAAGCTTACCTCTCCCTGGCGCTACAGTATCATCTTT 186

Db 87 TGGCCGCTTCAAGGAGACTTCAAGTACGCTGCTGCTGTCTCTAGCGCTGGTGTGC 146

Qy 187 ATCCTAGGCTGCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCGC 246

Db 147 GTGCTTGGGCTGTGTGAACGCCGTGGCGCTCTACATCTTCTTGGCGCCTCAAGACC 206

Qy 247 TGGAGCTGTGGCCACCACTATCTGTGAACCTGTGTGGCGGCGACCTGCTTTTATGTGCTA 306

Db 207 TGGAAATGGCTGCCACCACTATATGTTCACCTGGCTGTCTGTATGCATCATGTATGCGGCC 266

Qy 307 T---TGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGGTG 363

Db 267 TCCCTGGCGCTCTGTGTATTTACTACGCCCGCGGCGACCACTGGCCCTTTAGCAGCGGTG 326

Qy 364 CTCTCAAGCTGGTGCACCTTCTGTCTATATCAACCTTTTACGGCAGCATCTCTGTGTG 423

Db 327 CTCTCAAGCTGGTGGCTTCTCTTCTACACCAACCTTTTACTGCAGCATCTCTTCTCTC 386

Qy 424 ACCTGCATCTCTGTGCACAGTTCTTAGTGTGTGCCACCCACTGTGTTCGTGCGCTTAC 483

Db 387 ACCTGCATCAGCTGCACCGGTGTCTGGCGGCTTCTTACGACCTCTCGCTCCCTGGCGTGG 446

Qy 484 CGGACCCCGCAGCATGCTGTGGCGCACACGACCACTGGCGCCCTGTGTGCTCTCCAG 543

Db 447 GCGCGGCGCGCTACGCTCGCGGGGTGGCGGGCGGTGTGGGTGTGTGCTGGCCTTGC 506

Qy 544 CTGCTGCCCCACACTGGCCCTTCTCCACACGAGCTACATCAATGGCCAGATGCTGGTAT 603

Db 507 CAGGCCCGCGTGTCTACTTGTCTACACACGCGCGCGGGCGCGCTTAACCTGCCAC 566

Qy 604 GACATGACCAAGCAAGAAATTTTCATGGCTTTTGGCTAGCGCATAGTTCTGACATTG 663

Db 567 GACACCTCGCACCGAGCTCTTTCAGCGCTTCTGGCGCTACAGCTCATGTCTGGGC 626

Qy 664 TCTGGCTT--TCTTTCCCTCTCTTGTCTATTTTGGTGTGCTATTCATGATGTCTAGGAGC 721

Db 627 CTGCTCTTCGCGGTGCCCTTTGCCGCTCATCTTGTCTGTAGTGTCTATGCTCGGCGA 686

Qy 722 CTGATCAAGCCAGAGGAAACCTCATAGGACAGCAACACAGCCCGAGCCAGGTCCATC 781

Db 687 CTGCTAAAGCCAGCCCTACGGGACCTTCGGCGCGCTTGCCTAGGGCCAAAGCAGTCCG 746

Qy 782 CGGACCATCTACTGCTGTGGCCCTCTTACCCCTCTGTTTGTGCGCCCTTCATATCACT 841

Db 747 GGCACCATCGCGGTGTGCTGTGCTCTTTCGCGCTCTGCTNCTGCCATTCACAGTCCAC 806

Qy 842 CGCTCTCTTCTACCTCACCATCTGCTTCTTCTCTAGGACTGCCAGCTCTTGATGGCA 901

Db 807 CGCACCCTCTACTACTCTCTCG---TTCGCTGGACCTCAGCTGCCACACCTCAACGCC 863

Qy 902 GCCAGTGTGGCTTACAAGATATGGAGGCTCTGTGTGAGTGTGAGCAGTGTGCTCAACCCA 961

Db 864 ATCAACATGGCTTACAAGGTATACCGCGCGCTGGCGCAGTGTCTTAACAGTTGCTTGCACCC 923

Qy 962 GTCTGTACTTCTTTTCAAGG 982

Db 924 GTGCTACTCTCTGCTGGG 944

RESULT 11

AA542837/c  
ID AAS42837 standard; cDNA; 551 BP.

XX AC AAS42837;

XX DT 18-DEC-2001 (first entry)

XX DE Human G Protein-Coupled Receptor (GPCR) cDNA #32.

XX KW Human; G-protein coupled receptor; GPCR; mental disorder; schizophrenia; attention deficit disorder; anxiety; depression; bipolar disorder; ss;

XX KW neurological disorder; Huntington's disease; dementia; obesity; anorexia; metabolic disorder; Parkinson's disease; Tourette's syndrome; thrombosis; type 2 diabetes; cardiovascular disorder; myocardial infarction; cancer;

XX KW cardiomyopathy; atherosclerosis; human immunodeficiency virus; HIV; viral infection; immunostimulant; neuroleptic; nootropic; tranquilizer;

XX KW antidepressant; anorectic; PCR primer; gene therapy.

XX OS Homo sapiens.

XX PN WO200162797-A2.

XX PD 30-AUG-2001.

XX PF 23-FEB-2001; 2001WO-US05676.

XX PR 23-FEB-2000; 2000US-0184247.

XX PR 23-FEB-2000; 2000US-0184303.

XX PR 23-FEB-2000; 2000US-0184304.

XX PR 23-FEB-2000; 2000US-0184305.

XX PR 23-FEB-2000; 2000US-0184397.

XX PR 02-MAR-2000; 2000US-0186457.

XX PR 03-MAR-2000; 2000US-0186810.

XX PR 09-MAR-2000; 2000US-0188064.

XX PR 13-MAR-2000; 2000US-0188880.

XX PR 03-APR-2000; 2000US-0194344.

XX PR 23-JUN-2000; 2000US-0213861.

XX PR 11-JUL-2000; 2000US-0217369.

XX PR 11-JUL-2000; 2000US-0217370.

XX PR 14-JUL-2000; 2000US-0218337.

XX PR 20-JUL-2000; 2000US-0218492.

XX PA (PHRA ) PHARMACIA & UPJOHN CO.

XX PI Vogeli G, Wood LS, Parodi LA, Lind P;

XX DR WPI; 2001-570628/64.

XX DR P-PSDB; AAU25585.

XX PS New isolated nucleic acid encoding a new G-protein coupled receptor

XX PT polypeptide for detecting receptor modulators that can treat mental

XX PT disorders, such as schizophrenia, anxiety, depression, or obesity -

XX PS Claim 4; Page 82; 279pp; English.

XX CC Sequences AAS42806-AA542926 represent cDNA molecules and PCR primers for

XX CC cDNA molecules encoding human G-protein coupled receptor (GPCR)

XX CC polypeptides. The protein and DNA sequences of the invention can be used

CC viral infections caused by HIV and cancers.

XX SQ Sequence 551 BP; 143 A; 146 C; 130 G; 132 T; 0 other;

Query Match 20.5%; Score 222; DB 22; Length 551;

Best Local Similarity 100.0%; Pred. No. 2.1e-50;

Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTCCATTGCTTCCAGGGAAGCAGAGCGGAGCGGAGCGGAGCTCTG 60

DB 222 ATGCTGTCCATTGCTTCCAGGGAAGCAGAGCGGAGCGGAGCGGAGCTCTG 163

QY 61 CTCCTGGAGGAGCCCTCCCGGACATGGAGAAGGTGGACATGAATACACAGGAACAA 120

DB 162 CTCCTGGAGGAGCCCTCCCGGACATGGAGAAGGTGGACATGAATACACAGGAACAA 103

QY 121 GGTCTGTCCAGTTCTCAGAGAGTACAGCAAGTCTACTCTCCCTGGCCTACAGTATC 180

DB 102 GGTCTGTCCAGTTCTCAGAGAGTACAGCAAGTCTACTCTCCCTGGCCTACAGTATC 43

QY 181 ATCTTTATCTAGGGCTGCCACTAAATGGCACTGTCTTTGG 222

DB 42 ATCTTTATCTAGGGCTGCCACTAAATGGCACTGTCTTTGG 1

RESULT 12

AAQ88134

ID AAQ88134 standard; cDNA; 1842 BP.

XX AC AAQ88134;

XX DT 29-NOV-1995 (first entry)

XX DE Human P20 receptor gene.

XX KW Epithelial mucosa; mucus; cystic fibrosis; asthma; ss;

XX KW chronic bronchitis.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 57..1181

XX FT /\*tag= a

XX PN WO9510538-A.

XX PD 20-APR-1995.

XX PF 04-OCT-1994; 94WO-US11260.

XX PR 15-OCT-1993; 93US-0138137.

XX PA (UMOR ) UNIV MISSOURI.

XX PA (UYNC-) UNIV NORTH CAROLINA.

XX PI Boucher RC, Erb LJ, Harden TK, Lustig KD, Parr CE;

XX PI Sullivan DM, Turner JT, Weisman GA;

XX DR WPI; 1995-169967/22.

XX DR P-PSDB; AAR72457.

XX PT DNA encoding human P20 receptor and null cells expressing the

XX PT receptors - for stimulating or inhibiting growth of cultures of

XX PT mammalian cells, and for treating diseases of airway epithelial

XX PT mucosa, e.g. asthma.

XX PS Claim 3; Page 34; 47pp; English.

XX CC The sequence is that of the human P20 receptor gene. The gene may

XX CC be used to treat diseases of epithelial mucosal surfaces by

XX CC enhancing or inhibiting mucus prodn. The gene may be used to treat

XX CC cystic fibrosis, asthma and chronic bronchitis.

XX CC See also AAQ88135-40.

```
XX SQ Sequence 1842 BP; 364 A; 557 C; 528 G; 393 T; 0 other;
Query Match 18.7%; Score 202.2; DB 16; Length 1842;
Best Local Similarity 56.2%; Pred. No. 9e-45;
Matches 484; Conservative 0; Mismatches 363; Indels 14; Gaps 5;
QY 127 TGCCAGTCTCTCAGAGAAGTACAAGCAAGTCTACCTCTCCCTGGCTACAGTATCATCTTT 186
DB 129 TGCCGCTTCAAGAGGACTTGAAGTACGTGCTGCTGCTCTACGGCTGCTGCTG 188
QY 187 ATCTAGGCTGCCACTAAATGGCACTCTCTTTGGGCACTCTCTGGGCAACACCAAGCCG 246
DB 189 GTGCTTGGGCTGTGCTGAAGCCGCTGGGCGCTCTACATCTCTTCTTGGCGGCTCAAGACC 248
QY 247 TGGAGCTGTGGCACCACCTATCTGTGTGAACCTGATGTGGCGGACCTCTCTTTATCTGCTA 306
DB 249 TGGATGGCTGCACCAATATATGTTCACCTGGGCTGTCTGATGCACTGTATCGGCC 308
QY 307 T---TGCCCTTCTCATCATCACCTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTG 363
DB 309 TCCCTGCGCTGTCTTATTACTATACGCGCGCGGCGGCACTGGCCCTTCAGCACGGTG 368
QY 364 CTCTGCAAGCTGGTGCACCTCTCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTGCTG 423
DB 369 CTCTGCAAGCTGGTGGCTCTCTCTTCTTACACCACTTTTACTGCAGCATCTCTTCTCTC 428
QY 424 ACCTGCATCTCTGTGCACCAAGTCTCTAGTGTGTGCCACCACTGTGTGCTGCCCTAC 483
DB 429 ACCTGCATCAGCTGCACCGGTGCTCTGGCGCTCTTACGACCTCTCGCTCCCTGGCTGG 488
QY 484 CGGACCCGAGGATGCTGCTGGTGGGCAACGACCACTGGGCGCTGGTGTCTCTCAG 543
DB 489 GCGCGGCGCGCTAGCTGCGCGGTGGCGGCGGTGTGGG---TGTGGTGTGGCC 545
QY 544 CTGCTGCCACACTGGCTCTCTCCACAGGACTATCAATCAATGCGGCACTGATCTGGTAT 603
DB 546 TGCCAGGCGCGCTCTACTTTGTACACCAAGCGGCGGCGGCGGCTTACCTTGCAC 605
QY 604 GACATGACCAAGCAAGAAATTTGATCGGCTTTTGGCTACGGCATAGTTCTGACATTG 663
DB 606 GACACCTCGGCACCGAGCTCTTACAGCGCTTCGTGGCTACAGCTCAGTCACTGCTGGCC 665
QY 664 TCTGCTT--TCTTTCCTCTTGTGCTATTTTGGTGTGCTTACTGATGCTGAGGAGC 721
DB 666 CTGCTCTTTCGGGTGCGCTTTTGGCGTCACTCTGTCTGTACGTCTCATGCTCGGCGA 725
QY 722 CTGATCAAGCCAGAGGAACTCTATGAGGACAGCAACAGCCGAGGAGTCCATC 781
DB 726 CTGCTAAGCCAGGCTACGGGACCTTCGGGCGGCTGCTAGGCCCAAGTCCCGTG 785
QY 782 CGGACCACTCTACTGCTGTGTGGCTCTTACGCTCTCTTTTGGCGCTTCCATATCACT 841
DB 786 CGCACCATCGCGGTGTGCTGCTCTTTCGGCTCTCTCTCTCTCTCTCTCCATCCAGTCAAC 845
QY 842 CGCTCCTCTACTCACCACCTGCTTCTGCTTCTCAGGACTGCGGCTCTGATGAGCA 901
DB 846 CGCACCTCTACTACTCTCTCGCT--CGTGGACCTCAGTGTGCGACACGCTCAACGCC 902
QY 902 GCGAGTGTGGCTCAAGATATGAGGCTCTGTGTGAGTGTGAGGCTGCTCAACCCA 961
DB 903 ATCAACATGGGCTCAAGGT---TACCGGCTGGCGGCTGCTAAGCTTGCCTTGACCCC 959
QY 962 GTCTGTACTTCTTTTCAAGG 982
DB 960 GTGCTCTACTTCTGCTGGG 980
RESULT 13
AAD32937
ID AAD32937 standard; DNA; 2138 BP.
XX
AC AAD32937;
```

```
XX 01-JUL-2002 (first entry)
XX Mus musculus GPCR ATP-receptor P2U1 gene.
XX Mouse; G-protein coupled receptor; GPCR; transgenic animal; receptor;
XX genetic therapy; pharmacological; genetic disease; neuropsychological;
XX neurological; psychotic illness; gene; nephrotropic; gynaecological;
XX psychostimulant; ATP-receptor P2U1; ds.
XX Mus musculus.
XX Key Location/Qualifiers
XX CDS 270..1391
XX FT /*tag= a
XX FT /product= "ATP-receptor P2U1 protein"
XX FT misc_feature 1569..1570
XX FT /*tag= b
XX FT /note= "Illegible in the specification"
XX WO200203793-A2.
XX 17-JAN-2002.
XX 10-JUL-2001; 2001WO-US21923.
XX 10-JUL-2000; 2000US-217058P.
XX 10-JUL-2000; 2000US-217179P.
XX 10-JUL-2000; 2000US-217223P.
XX 10-JUL-2000; 2000US-217253P.
XX 10-JUL-2000; 2000US-217255P.
XX 10-JUL-2000; 2000US-217256P.
XX 10-JUL-2000; 2000US-217257P.
XX 11-JUL-2000; 2000US-217347P.
XX 11-JUL-2000; 2000US-217629P.
XX 12-JUL-2000; 2000US-217537P.
XX 12-JUL-2000; 2000US-218069P.
XX 12-JUL-2000; 2000US-218074P.
XX 12-JUL-2000; 2000US-218358P.
XX 27-JUL-2000; 2000US-221483P.
XX 07-AUG-2000; 2000US-223120P.
XX 07-AUG-2000; 2000US-223122P.
XX 26-OCT-2000; 2000US-243958P.
XX 15-NOV-2000; 2000US-249408P.
XX 20-NOV-2000; 2000US-252299P.
XX 16-JAN-2001; 2001US-262113P.
XX 16-JAN-2001; 2001US-262205P.
XX (DELT-) DELTAGEN INC.
XX Allen KD, Brennan TJ;
XX WPI: 2002-164574/21.
XX P-PSDB: AAE20604.
XX Novel non-human transgenic animal, especially transgenic mice useful
XX for identifying an agent that modulates expression or function of
XX target gene, comprises disruptions in target G protein-coupled receptor
XX gene.
XX Disclosure: Fig 19; 101pp; English.
XX The invention relates to a non-human transgenic animal having targeted
XX G-protein coupled receptor (GPCR) gene disruptions in melanocortin-3 gene
XX 5-HT-2B gene, chemokine receptor 9A gene, glucocorticoid-induced receptor
XX gene, orphan GPR10 (UHR)-1 gene, orphan GPR14 gene, orphan GPR15 gene,
XX beta chemokine receptor (E01) gene, endothelial differentiation GPCR3
XX (EDG3) gene, ATP receptor P2U1 gene or adenosine 3 receptor gene. The
XX transgenic animal is useful for identifying an agent that modulates the
XX expression or function of the target gene, for identifying an agent that
XX ameliorates a phenotype associated with a disruption in the target gene.
XX A transgenic construct is useful for producing a transgenic animal,
XX preferably a transgenic mouse. The transgenic animal is useful for
```

CC testing the efficacy of proposed genetic and pharmacological therapies  
CC for human genetic diseases, such as neurological, neuropsychological or  
CC psychotic illnesses. The transgenic animal is also useful as models for  
CC diseases, disorders or conditions associated with phenotypes relating to  
CC a disruption in a target, and to identify pharmaceuticals, therapies,  
CC drugs and interventions which may be effective in treating a disease or  
CC other phenotypic characteristics of the animal. An agent which modulates  
CC the expression of the target gene is useful as a therapeutic for treating  
CC conditions associated with a disruption of the target gene. The present  
CC sequence is mouse GPCR ATP-receptor P2U1 gene.  
XX  
SQ

Sequence 2138 BP; 423 A; 623 C; 606 G; 480 T; 6 other;	
Query Match	18.7%; Score 202.2; DB 24; Length 2138;
Best Local Similarity	54.6%; Pred. No. 9.6e-45;
Matches 470; Conservative 0; Mismatches 383; Indels 8; Gaps 3;	
QY 127 TGGCAGTCTTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186	
DB 342 TGTCTGCTTCAACGAGGACTTCAAGTACGTGCTGTTGCCGCTGTCTATGGCGTGTGTC 401	
QY 187 ATCTAGGCTGCCACTAATGCACTGCTGTGGCACTCTCTGGGCCCAACCAAGCGC 246	
DB 402 GTCTCGGGTGTGCTGACGCTGCGTCTCTATCTCTTCTATGCGCGCTCAAAACC 461	
QY 247 TGGAGCTGTGCCACCACTATCTGTGGAACCTGATGGTGGCCGCACTGCTTTATGTGCTA 306	
DB 462 TGAACGCTCCACCACTACATGTTTACCTGGCAGTTTCGCACTCTCTACGAGCGC 521	
QY 307 T----TGCCCTTCTCATCATACCTACTACTAGATGACAGGTGGCCCTTCGGGGAGCTG 363	
DB 522 TCCCTCGCGCTGTGTTTATTATACCGCGGGGTGACCACTGGCCATTTAGCAGCGTG 581	
QY 364 CTCTGAAGCTGTGCATCTCTGTTCTATATCAACTTTACGGCAGCATCTCTGCTGCTG 423	
DB 582 CTCTGAAGCTGTGCTGTTTCTCTTCTACCAACCTCTACTGACGACATCTCTCTCCCT 641	
QY 424 ACCTGATCTCTGTGACCACTAGTTAGTGTGTCACCACTGTTGCTGCGCTCCCTAC 483	
DB 642 ACCTGATCAGCTGACCGGTGCTGGAGTCTCTGCGCTCTGCACTCCCTGCGTGG 701	
QY 484 CGGACCGGAGGATGCTGCTGGTGGGACAGCAGCACTGGCCCTGGTGGTCTCCAG 543	
DB 702 GGCGGGCGCGTTATGCCCGCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 761	
QY 544 CTGCTGCCACACCTGCTCTCCCAACAGGACTACATCAATGGCCAGATGATCTGGTAT 603	
DB 762 CAGGACCGCTGCTCTACTTCGTACCAACAGCAGCGTGGGGGAACCGGATCATCTGCCAT 821	
QY 604 GACATGACCAAGCAAGAGAAATTTGATCGGCTTTTTCCTACGGCATAGTTCTG--ACAT 661	
DB 822 GACACCTCGCGCGAGAGCTCTTTAGCCATTTTGTGGCTTACAGCTCCGTCATGCTGGT 881	
QY 662 TGTCTGGCTTCTTCTGCTGCTGCTATTTGGTGTGCTATTTACTATGATGGTCAAGAGC 721	
DB 882 CTGCTTTTGTGCTGCGCTTTTCCGTAATCTGCTGTGTTACGCTGTATGGCCAGCGG 941	
QY 722 CTGATCAAGCAGAGAGAACCTCATGAGCAGCAACACCGCCAGCCAGCTCCATC 781	
DB 942 CTGCTCAACCGGCTTATGGGACCAAGAGGTCTGCTCGGGGCAAGCAAGTCTGTG 1001	
QY 782 CGGACCATCTACTGTGTGGCCCTCTTCCACCTCTGTTTGTGGCCCTTCCCATATCACT 841	
DB 1002 CGACCATTTGCTTGTGCTGCTGCTCTGCGCTCTGCTTTCTGCTTTCCAGCTCAG 1061	
QY 842 CGCTCTCTTCACTACCATCTGCTTCTGCTTTCTGAGGAGTGGCAGCTCTTTGATGGCA 901	
DB 1062 CGCACCTCTACTACTCTCTCCGAT---CACTTGACCTCAGCTGCCACACCTCAAGCC 1118	
QY 902 GCCAGTGTGGCTTACAGATATGGAGCCCTCTGTTGAGTGTGAGCAGCTGCCCTCAACCCA 961	
DB 1119 ATCAACATGGCATATAGATCACCGCGCGCTGGCGAGCGCCCAACAGTCTCTTTGACCG 1178	

QY 962 GTCCTGCTACTTCTTTTCAAGG 982	
DB 1179 GTACTCTACTTCTGCGAGG 1199	
RESULT 14	
AAD32944	
ID AAD32944 standard; DNA; 2138 BP.	
XX AAD32944;	
AC AAD32944;	
XX 01-JUL-2002 (first entry)	
XX Mus musculus GPCR ATP-receptor P2U1 gene targeting construct.	
DE Mouse; G-protein coupled receptor; GPCR; transgenic animal; receptor;	
XX genetic therapy; pharmacological; genetic disease; neuropsychological;	
KW neurological; psychotic illness; nephrotropic; gynaecological;	
KW psychostimulant; ATP-receptor P2U1; ds.	
XX Mus musculus.	
OS	
XX Key Location/Qualifiers	
FH misc_feature 286..350	
FT /tag= a	
FT /note= "Sequence flanking Neo insert in targeting	
FT construct"	
FT misc_feature 351..636	
FT /tag= b	
FT /note= "Sequence deleted in targeting construct"	
FT misc_feature 637..941	
FT /tag= c	
FT /note= "Sequence flanking Neo insert in targeting	
FT construct"	
XX WO200203793-A2.	
PD 17-JAN-2002.	
XX 10-JUL-2001; 2001WO-US21923.	
PF	
XX 10-JUL-2000; 2000US-217058P.	
PR 10-JUL-2000; 2000US-217179P.	
PR 10-JUL-2000; 2000US-217223P.	
PR 10-JUL-2000; 2000US-217253P.	
PR 10-JUL-2000; 2000US-217255P.	
PR 10-JUL-2000; 2000US-217256P.	
PR 10-JUL-2000; 2000US-217257P.	
PR 11-JUL-2000; 2000US-217347P.	
PR 11-JUL-2000; 2000US-217629P.	
PR 12-JUL-2000; 2000US-217537P.	
PR 12-JUL-2000; 2000US-218069P.	
PR 12-JUL-2000; 2000US-218074P.	
PR 12-JUL-2000; 2000US-218358P.	
PR 27-JUL-2000; 2000US-221483P.	
PR 07-AUG-2000; 2000US-223120P.	
PR 26-OCT-2000; 2000US-223122P.	
PR 26-OCT-2000; 2000US-243958P.	
PR 15-NOV-2000; 2000US-249408P.	
PR 20-NOV-2000; 2000US-252299P.	
PR 16-JAN-2001; 2001US-262113P.	
PR 16-JAN-2001; 2001US-262205P.	
XX (DELT-) DELTAGEN INC.	
XX Allen KD, Brennan TJ;	
PI WPI; 2002-164574/21.	
XX Novel non-human transgenic animal, especially transgenic mice useful	
PT for identifying an agent that modulates expression or function of	
PT target gene, comprises disruptions in target G protein-coupled receptor	
PT gene	









GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 21:21:18 ; Search time 62 Seconds  
(without alignments)  
5352.004 Million cells

Title: US-10-023-586B-1

Perfect score: 1082  
Sequence: 1 atgctgtccattttgtcttc.....ccagggttgacacagatctgg 1082

scoring table: IDENTITY\_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

```

Post processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```
Database :
Issued_Patents_NA.*
1: /cgn2_6/pdata/1/1na/5A_COMB.seq.*
2: /cgn2_6/pdata/1/1na/5B_COMB.seq.*
3: /cgn2_6/pdata/1/1na/6A_COMB.seq.*
4: /cgn2_6/pdata/1/1na/6B_COMB.seq.*
5: /cgn2_6/pdata/1/1na/pTUS_COMB.seq.*
6: /cgn2_6/pdata/1/1na/backfiles1.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

result No.	Query			DB	ID	Description
	Score	Match	Length			
1	203.8	18.8	1842	1	US-08-442-134A-1	Sequence 1, Appli
2	203.8	18.8	1842	1	US-08-444-581B-1	Sequence 1, Appli
3	203.8	18.8	1842	1	US-08-446-088A-1	Sequence 1, Appli
4	164.8	15.2	984	3	US-08-513-974B-57	Sequence 57, Appl
5	164.8	15.2	1023	3	US-08-513-974B-379	Sequence 379, Appl
6	156.8	14.5	984	3	US-08-459-046-1	Sequence 1, Appli
7	148.8	13.8	984	3	US-08-513-974B-41	Sequence 41, Appl
8	148.8	13.8	1020	3	US-08-513-974B-370	Sequence 370, Appl
9	92	8.5	1475	1	US-08-097-938-1	Sequence 1, Appli
10	92	8.5	1475	1	US-08-476-000-1	Sequence 1, Appli
11	92	8.5	1475	1	US-08-472-840-1	Sequence 1, Appli
12	92	8.5	1475	2	US-08-476-976-1	Sequence 1, Appli
13	92	8.5	1475	3	US-08-474-410-1	Sequence 1, Appli
14	92	8.5	1475	4	US-08-486-673B-1	Sequence 1, Appli
15	92	8.5	2732	1	US-08-476-000-60	Sequence 60, Appl
16	92	8.5	2732	2	US-08-472-840-60	Sequence 60, Appl
17	92	8.5	2732	2	US-08-476-976-60	Sequence 60, Appl
18	92	8.5	2732	3	US-08-474-410-60	Sequence 60, Appl
19	92	8.5	2732	4	US-08-486-673B-60	Sequence 60, Appl
20	91.4	8.4	1205	1	US-08-417-103-13	Sequence 13, Appl
21	91.4	8.4	1634	1	US-07-816-283-1	Sequence 1, Appli
22	91.4	8.4	1634	1	US-08-417-103-1	Sequence 1, Appli
23	88.4	8.2	1901	3	US-08-153-848-43	Sequence 43, Appl
24	88.4	8.2	1901	3	US-09-299-843A-43	Sequence 43, Appl
25	88.4	8.2	1901	4	US-09-088-337B-43	Sequence 43, Appl
26	88.4	8.2	1901	5	PCF-US93-11153-43	Sequence 43, Appl
27	88.4	8.2	2453	5	PCF-US95-07180-1	Sequence 1, Appli

Sequence 6,	Appli
Sequence 5,	Appli
Sequence 5,	Appli
Sequence 5,	Appli
Sequence 5,	Appli
Sequence 5,	Appli
Sequence 5,	Appli
Sequence 5,	Appli
Sequence 5,	Appli
Sequence 5,	Appli
Sequence 3,	Appli
Sequence 7,	Appli
Sequence 1,	Appli
Sequence 3,	Appli
Sequence 3,	Appli
Sequence 1,	Appli
Sequence 7,	Appli
Sequence 7,	Appli
Sequence 3,	Appli

## ALIGNMENTS

```

RESULT 1
US-08-442-134A-1
; Sequence 1, Application US/08442134A
; Patent No. 5596088
; GENERAL INFORMATION:
; APPLICANT: Boucher, Richard C.
; APPLICANT: Weisman, Gary A.
; APPLICANT: Turner, John T.
; APPLICANT: Harden, Thomas K.
; APPLICANT: Parr, Claude E.
; APPLICANT: Sullivan, Daniel M.
; APPLICANT: Erb, Laura
; APPLICANT: Lustig, Kevin D.
; TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and
; TITLE OF INVENTION: Null Cells Expressing P2U Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5596088th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/442,134A
; FILING DATE: 16-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 5470-71A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-420-2200
; TELEFAX: 919-881-3175
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1842 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..1181
US-08-442-134A-1

```

Query Match 18.8%; Score 203.8; DB 1; Length 1842;  
Best Local Similarity 56.3%; Pred. No. 1.7e-46;  
Matches 485; Conservative 0; Mismatches 362; Indels 14; Gaps 5;

Qy 127 TGCAGTTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186  
Db 129 TGCAGTTCAGAGAGTTCAGAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186  
Qy 187 ATCTAGGCTGCGCCAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 246  
Db 189 GTGCTGGGCTGTGTCTGAACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 248  
Qy 247 TGGAGCTGTGCGCCACCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 306  
Db 249 TGGAGCTGTGCGCCACCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 308  
Qy 307 T---TGCCCTTCTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 363  
Db 309 TCCCTGGCGCTGCTGGTCTATTACTACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCG 368  
Qy 364 CTCTGCAAGCTGGTGCACCTTCTCTGTATATATATATATATATATATATATATATAT 423  
Db 369 CTCTGCAAGCTGGTGGGCTTCTCTTCTTACACCACTTCTTCTTCTTCTTCTTCTTCT 428  
Qy 424 ACCTGCATCTCTGTGCACCACTTCTCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 483  
Db 429 ACCTGCATCTCTGTGCACCACTTCTCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 488  
Qy 484 CGGACCCGAGGAGTCTGT 543  
Db 489 GCGCGGCGCGGCTAGCTGT 545  
Qy 544 CTGCTGCGCCACCTGCTGT 603  
Db 546 TCCAGGCGCGGCTGCTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 605  
Qy 604 GACATGACGCGCAAGAAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 663  
Db 606 GACACCTCGGCAACCGAGCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 665  
Qy 664 TCTGGCTT-TCTTTCCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 721  
Db 666 CTGCTCTTGGGCTGCTTGT 725  
Qy 722 CTGATCAAGCGCAGAGAGAACTCTATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781  
Db 726 CTGCTAAAGCCAGCTACGGGACCTTCGGGCGGCTTCGGGCGGCGGCGGCGGCGGCGG 785  
Qy 782 CGGACCATCTACTGT 841  
Db 786 CGCACCATCGCGGT 845  
Qy 842 CGCTCTTCTACTCACCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 901  
Db 846 CGCACCTCTACTACTCTCTCGGT--CGTGGACCTCAGCTGCGCACCACTCAACGCC 902  
Qy 902 GCCAGTGTGGCTACAAGATATGAGGCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 961  
Db 903 ATCAACATGGCTACAAGT---TACCGGCTGGCGCACTGTCTAAGCTGTGCTTGAACCCC 959  
Qy 962 GTCTGTACTTCTTCAAGG 982  
Db 960 GTGCTACTTCTTGGCTGGG 980

## RESULT 2

US-08-444-581B-1  
; Sequence 1, Application US/08444581B  
; Patent No. 5607836  
; GENERAL INFORMATION:  
; APPLICANT: Boucher, Richard C.  
; APPLICANT: Weisman, Gary A.

APPLICANT: Turner, John T.  
APPLICANT: Harden, Thomas K.  
APPLICANT: Parr, Claude E.  
APPLICANT: Sullivan, Daniel M.  
APPLICANT: Erb, Laura  
APPLICANT: Lustig, Kevin D.  
TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and  
TITLE OF INVENTION: Nucleic Acids Expressing P2U Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bell, Seitzer, Park & Gibson  
STREET: Post Office Drawer 34009  
CITY: Charlotte  
STATE: No. 5607836th Carolina  
COUNTRY: USA  
ZIP: 28234  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,581B  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/442,134  
FILING DATE: 16-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Sibley, Kenneth D.  
REGISTRATION NUMBER: 31,665  
REFERENCE/DOCKET NUMBER: 5470-71A  
TELEPHONE: 919-420-2200  
TELEFAX: 919-881-3175  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1842 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 57..1181  
US-08-444-581B-1

Query Match 18.8%; Score 203.8; DB 1; Length 1842;  
Best Local Similarity 56.3%; Pred. No. 1.7e-46;  
Matches 485; Conservative 0; Mismatches 362; Indels 14; Gaps 5;

Qy 127 TGCAGTTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186  
Db 129 TGCAGTTCAGAGAGTTCAGAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTT 186  
Qy 187 ATCTAGGCTGCGCCAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 246  
Db 189 GTGCTGGGCTGTGTCTGAACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 248  
Qy 247 TGGAGCTGTGCGCCACCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 306  
Db 249 TGGAGCTGTGCGCCACCACTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 308  
Qy 307 T---TGCCCTTCTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 363  
Db 309 TCCCTGGCGCTGCTGGTCTATTACTACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCG 368  
Qy 364 CTCTGCAAGCTGGTGCACCTTCTCTGTATATATATATATATATATATATATATATAT 423  
Db 369 CTCTGCAAGCTGGTGGGCTTCTCTTCTTACACCACTTCTTCTTCTTCTTCTTCTTCT 428  
Qy 424 ACCTGCATCTCTGTGCACCACTTCTCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 483  
Db 429 ACCTGCATCTCTGTGCACCACTTCTCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 488

Db 429 ACTGCATCAGCGTGCACCGGTGCTGGGGCTTTACGACCTCTCGGCTCCCTGGCGCTGG 488  
QY 484 CGGACCCGAGGAGATGCTGGTGGGACACGACACCTGGGCTGGTGGTCTCCAG 543  
Db 489 GCGCGGCGGCTACGCTCGCGGTGGCGGGCGGTGGG---TGTGGTGGTGGCC 545  
QY 544 CTGCTGCCACACTGCGCTTCTCCACACGAGTATACATCAATGAGGCGAGATGATGAT 603  
Db 546 TGCCAGGCGCGGTGCTCTACTTTGTACACGAGCGCGGGCGCTAACCTGCCAC 605  
QY 604 GACATGACGAGCAAGAGAAATTTGATCGCTTTTGGCTTGGCTTGGCTTGGCTTGGCT 663  
Db 606 GACACCTGGCAGCGGCTCTACGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 665  
QY 664 TCTGGCTT---TCTTTCCCTTGGTCAATTTGGTGGCTTGGCTTGGCTTGGCTTGGCT 721  
666 CTGCTCTTGGCGGTGCGCTTGGCGGTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 725  
QY 722 CTGATCAAGCAGGAGAACTCTATGAGGAGGCAACAGCGCGGCTAGGCGGCTAGGCGG 781  
Db 726 CTGCTAAAGCAGGAGAACTCTATGAGGAGGCAACAGCGCGGCTAGGCGGCTAGGCGG 785  
QY 782 CGGACCATCTACTGCTGGCTTCTCCACACGAGTATACATCAATGAGGCGAGATGATGAT 841  
Db 786 CGGACCATCTACTGCTGGCTTCTCCACACGAGTATACATCAATGAGGCGAGATGATGAT 845  
QY 842 CGCTCCTCTACTGCTGGCTTCTCCACACGAGTATACATCAATGAGGCGAGATGATGAT 901  
Db 846 CGCACCCTCTACTGCTGGCTTCTCCACACGAGTATACATCAATGAGGCGAGATGATGAT 902  
QY 902 GCGAGTGGCTTACAGATGAGGCGCTTGGTGGCTTGGTGGCTTGGTGGCTTGGTGGCT 961  
Db 903 ATCAACATGCGCTTACAGGCT---TACCGGCTGGCGAGTGAACAGTGGCTTGGCTTGG 959  
QY 962 GTCTCTACTTCTTCAAGG 982  
Db 960 GTGCTCTACTTCTGGCTGGG 980

RESULT 3

US-08-446-088A-1

; Sequence 1, Application US/08446088A

; Patent No. 5691156

GENERAL INFORMATION:

APPLICANT: Boucher, Richard C.

APPLICANT: Weisman, Gary A.

APPLICANT: Turner, John T.

APPLICANT: Harden, Thomas K.

APPLICANT: Parr, Claude E.

APPLICANT: Sullivan, Daniel M.

APPLICANT: Erb, Laura

APPLICANT: Lustig, Kevin D.

TITLE OF INVENTION: DNA Encoding the Human P2U Receptor and

TITLE OF INVENTION: Null Cells Expressing P2U Receptors

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Bell, Seltzer, Park & Gibson

STREET: Post Office Drawer 34009

CITY: Charlotte

STATE: No. 5691156th Carolina

COUNTRY: USA

ZIP: 28234

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/446,088A

FILING DATE: 19-MAY-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kenneth D. Sibley

; REGISTRATION NUMBER: 31,665  
; REFERENCE/DOCKET NUMBER: 5470-71C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 919-420-2200  
; TELEFAX: 919-881-3175  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1842 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 57..1181  
; US-08-446-088A-1

Query Match 18.8%; Score 203.8; DB 1; Length 1842;  
Best Local Similarity 56.3%; Pred. No. 1.7e-46;  
Matches 485; Conservative 0; Mismatches 362; Indels 14; Gaps 5;  
QY 127 TGCCAGTCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATCATCTTT 186  
Db 129 TGCCGCTTCAAGAGAGACTTCAAGTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 188  
QY 187 ATCCTAGGCTGCCACTAAATGGCACTGCTTGTGGCACTCTCTGGGCGGCAAAACAGCGC 246  
Db 189 GTGCTGGGCTGTGCTGAACCGCTGGGCTCTACATCTTCTGTGGCGCTCAAGACC 248  
QY 247 TGGAGTGTGGCAACCACTATCTGGTGAACCTGATGGTGGCGGACCTGCTTTATGCTA 306  
Db 249 TGGAAATGCGCTCCACCACATATATGTTCCACCTGGCTGTGTCTGATGCACTGTATGCG 308  
QY 307 T---TGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCGCTTCGGGGAGCTG 363  
Db 309 TCCTGCGCGTGTGCTGCTATTACTACGCGCGGCGGCGGCGGCTTCAAGCGGTG 368  
QY 364 CTCTGAAGCTGTGCACTTCTCTTCTATATCAAGCTTTTACGGCAGCATCTCTGCTGCTG 423  
Db 369 CTCTGAAGCTGTGCGCTTCTCTTCTACACCACTTTTACGAGCATCTCTTCTCTC 428  
QY 424 ACCTGATCTCTGTGCAAGTCTCTAGTGTGTGGCCACCACTGTGTGGCTGCGCTAC 483  
Db 429 ACCTGATCAGGCTGCACGGGTGCTGTGGGCTCTTACGACCTCTGCGCTCCCTGCGCTG 488  
QY 484 CGGACCCGAGGAGTCTGGTGGCGCACGACACCACTGGCGCTGTGGTGGCTTCCAG 543  
Db 489 GCGCGGCGCGCTTACCTCGCGGGTGGCGGCGGCGGTGTGGG---TGTGGTGGTGGCC 545  
QY 544 CTGCTGCCACACTTGGCTTCTCCACAGGAGTACATCAATGGCCAGATGATCTGCTAT 603  
Db 546 TGCCAGGCGCGCTGCTCTACTTGTACACAGCGCGCGCGGCTTAACCTGCCAC 605  
QY 604 GACATGACGAGCAAGAGAAATTTGATCGGCTTTTGGCTTACGCGCATAGTCTTCAATG 663  
Db 606 GACACCTCGGCAACCGAGCTCTCAGCGCTTCTGGGCTTACAGCTCAGTCACTGCGGCG 665  
QY 664 TCTGGCTT---TCTTTCCCTTGGTCAATTTGGTGTGCTATTTCATCTGATGGTCAAG 721  
Db 666 CTGCTCTTCCGCGTGGCTTGGCGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 725  
QY 722 CTGATCAAGCAGGAGAACTCTATGAGGAGGCAACAGCGCGGCTAGGCGGCTAGGCGG 781  
Db 726 CTGCTAAAGCAGGAGAACTCTACGGGAGCTCGGGGCGCTGCTAGGCGGCAAGTCCG 785  
QY 782 CGGACCATCTACTGCTGGTGTGGCTTCTTACCCCTCTTACCCCTCTGTTTGTGGCTTCC 841  
Db 786 CGCACCATCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 845  
QY 842 CGCTCCTTCTACTGCTACCATCTGCTTTCTGCTTCTTCTGCTGCTGCTGCTGCTGCTG 901  
Db 846 CGCACCCTCTACTGCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 902



Db	831	GGCCTTTGCAGCGCCTACAAAGGCACGCGCGCTTTGCCAGTGCACACGCTGCTGGA	890
Qy	958	CCCAGTCTCTGTA	971
Db	891	CCCCATCCTCTTCT	904

## 5. RESULT

US-08-513-974B-379  
; Sequence 379, Application US/08513974B  
; Patent NO. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ontaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B  
; FILING DATE: 14-SEP-1995  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP95/01599  
; FILING DATE: 10-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-093989  
; FILING DATE: 19-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-057186  
; FILING DATE: 16-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-007177  
; FILING DATE: 20-JAN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-326611  
; FILING DATE: 28-DEC-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-270017  
; FILING DATE: 02-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236357  
; FILING DATE: 30-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-236356  
; FILING DATE: 30-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189274  
; FILING DATE: 11-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189273  
; FILING DATE: 11-AUG-1945  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-189272  
; FILING DATE: 11-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Resnick, David S.  
; REGISTRATION NUMBER: 34,235

```

: REFERENCE/DOCKET NUMBER: 45753
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO.: 379;
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1023 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 37..1020
: US-08-513-974B-379

Query Match      15.2%; Score 164.8; DB 3; Length 1023;
Best Local Similarity 52.0%; Pred. No. 6.7e-36;
Matches 444; Conservative 0; Mismatches 402; Indels 8; Gaps

Qy 126 CTGCCAGTTCTCAGAGAAGTACAAAGTCTACCTCTCCTCGCCTACAGTATCATCTT 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 87 CTGTGTCTACCGGAGAACTTCAAGCAACTGCTGCGCCACCTGTGTATTGCGGTGCT 146

Qy 186 TATCCTAGGCGTGCACATAAATGGCACTGCTTGTGGCACTCTCTGGGGCCAAACCAAGG 245
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 GCGGCTGGGCTGCGGTGAACATCTGTGTCAATTAGCCAGATCTGCACGTCCGCGCGGC 206

Qy 246 CTGGAGCTGTGCCACCACCTATCTGGTGAACCTGATGGTGGCGACCTGCTTTATGTGCT 305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 CCTGACCCTGACGGCGGTGACACCTAAACCTTGTCTGGCTGACCTGCTATATGCTTG 266

Qy 306 AT---TGCCCTTCTCATCATCACTACTCACTAGATGACAGGTGGCCCTTCGGGGAGCT 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 CTCCTCGCCCTGCTCATCTACAACATATGCCAAGGTGATCACTGGCCCTTTGGCGACTT 326

Qy 363 GCTCTGAAGCTGGTGACATTCTCTGTTCTATATCAACCTTTTAGGCAGAGATCTGCTGCT 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 327 CGCCTGCGGCTGGTGGCGCTTCCTCTTCTATGCCAACCTGCACGGCAGCATCTCTTCCT 386

Qy 423 GACCTGCATCTCTGTGCACAGTTCTCTAGGTGTGTCACCCCACTGTGTCGCTGCCCTA 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 387 CACCTGCATCAGCTTTCAGCGCTACCTGGGCATCTGCCACCCGCTGGCCCTGGCACAA 446

Qy 483 CC---GGACCCAGGCGATGCCTGGCTGGGCACACAGCACCTCGGCCCTGGTGGCTCT 539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 447 ACGTGGGGCGCGCGCTGCTGGCTAGTGTGTGAACGCTGCTGGCTGGCGTGACAC 506

Qy 540 CCAGCTGTCGCCACACTGGCTTCTCCACACGGACTACATCAATGGCCAGATGATCTG 599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 507 CCAGTGGCTGCCACAGCCATCTTCGCTGTCACAGGCGATCCAGCGTAAACGCGACTGCT 566

Qy 600 GTATGACATGACCAAGCAATTTTGATGGCTTTTGTGCTACGGATAGCTTCTGCAC 659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 567 CTATGACCTCAGCCGCTGCCCTGGCCACCACCTATATGCCCTATGGCATGGCTCTCAC 626

Qy 660 ATTGCTGCTTTTCT--TTCCCTCTTGGTCATTTTGGTGTGCTATTCTAGTATGCTCAG 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 627 TGTATCGGCTTCTGCTGCCCTTTGCTGCCCTGCTGGCCTGCTACTGCTCTCTGGCCTG 686

Qy 718 GAGCCTGATCAAGCCAGGAGAACCTCATGAGACAGGCAACACAGCCGAGCCAGGTC 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 687 CCGCCTGTGGCGCAGGATGGCCCGGACAGACGCTGTGGCCACGAGCGGCTGGCAAGGC 746

Qy 778 CATCGGACCATCCTACTGGTGTGTCCTTTCACCCCTCTGTTTGTGCCCTTCCATAT 837
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 747 GGCCCGCATGGCGGTGTGGTGGCTTGTGCTTTGGCATCAGCTTCTCTGCTTTTCACAT 806

Qy 838 CACTCGCTCCTTCTACCTCACCATCTCTTTTCTGCTTTCTCAGGACTGCCAGCTCTTGAT 897
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 807 CACCAAGACACGCTTACCTGGCAGTGGGCTCGACGCGGGGCTCCCTGCTACTATTGA 866

Qy 898 GGCAGCCAGTGTGGCCTACAAGATATGAGAGCCCTCTGGTGAGTGTGAGCAGCTGCCCTAA 957
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 867 GGCCTTTGACGGCCTACAAAGGACGCGCGCTTTGCCAGTGCCAAACAGCGTGCTGGA 926  
Qy 958 CCCAGTCTGTACT 971  
Db 927 CCCATCCTCTTCT 940

RESULT 6  
US-08-459-046-1  
; Sequence 1, Application US/08459046  
; Patent No. 6008039  
; GENERAL INFORMATION:  
; APPLICANT: Coleman, Roger  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Stuart, Susan G.  
; TITLE OF INVENTION: A No. 6008039el Human Purinergic P2U Receptor  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3330 Hillview Avenue  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Luther, Barbara J.  
; REGISTRATION NUMBER: 33954  
; REFERENCE/DOCKET NUMBER: PF-0038 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-852-0195  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 984 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; IMMEDIATE SOURCE:  
; LIBRARY: Placenta  
; CLONE: 179696  
US-08-459-046-1

Very Match 14.5% Score 156.8; DB 3; Length 984;  
Best Local Similarity 51.4%; Pred. No. 1e-33;  
Matches 439; Conservative 0; Mismatches 407; Indels 8; Gaps 3;

Qy 126 CTGCAGTCTTCAGAGAGTACAAGCAAGTCTACCTCTCCCTGGCCTACAGTATCATCTT 185  
Db 51 CTGTGTTCTACCGCAGAGAACTTCAAGCAACTGCTGCTCCACCTGTGTATTCGGCGGTGCT 110  
Qy 186 TATCCTAGGCGTCCCACTAAATGGCACTGTCTTGTGGCACTCTCTGGGCGCAACCAAGCG 245  
Db 111 GCGCGCTGCGCTCCGCTGGAACATCTGTGTCAATTACCCAGATCTGCACGTCCCGCGGGC 170  
Qy 246 CTGAGAGTGTGCCACCACTATCTGGTGAACCTGATGTGGCGCACCTGCTTTATGTGCT 305  
Db 171 CTGACCGGACGCGCGGTGACACCCCTAACCCTTGCTGCTGCTGCTGCTATATGCGTG 230  
Qy 306 AT---TGCCCTTCTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 362  
Db 231 CTCCTGCGCCCTGCTCATCTACAACTATGCGCAAGGTGATCATCTGGCCCTTTGGCGACTT 290

Qy 363 GCTCTGCAAGTGGTGTGCACTTCCTCTGTCTATATCAACCTTTACGGCAGCATCCTCTCTGCT 422  
Db 291 CGCCTGCGCGCTGGTGTGCGCTTCTCTATATGCAACCTGCACGGAGGATCCTCTTCTCT 350  
Qy 423 GACCTGCATCTCTGTGCAACAGTTCCCTAGGTGTGTGCGCACCACTGTGTCTGCTGCCCTA 482  
Db 351 CACCTGCATCAGCTTCCAGCGCTACCTGGGCATCTGCGCACCGCTGGCCCTTGGCCCA 410  
Qy 483 CC---GGACCCGCGAGCATGCTGGTGGCACCAGCAGCACCTGGCGCTGGTGTGCTCT 539  
Db 411 ACGTGGGGCGCGCGGCTGCTGGCTAGTGTGTGTACCGCTGTGGCTGGCGTGACAC 470  
Qy 540 CCAGTGTGTCGCCACACATGGGCTTCTCCACAGGAGTACATCAATGGCCAGATGATCTG 599  
Db 471 CCAGTGTGTCGCCACAGCATCTTCTGCTGCCAGCATCCAGCGTAAACCGCACTGTCTG 530  
Qy 600 GTATGACATGACGAGCCAAAGAAATTTTGTATGCGGCTTTTGTGCTACGCGCATAGTTCTGAC 659  
Db 531 TTATGACCTCAGCCGCGCTGCGCTGGCCACCCACTATATGCGCTATGGGATGGCTCTCAC 590  
Qy 660 ATTGTCTGGCTTTCT--TTCCCTCTCTTGGTTCATTTTGGTGTGCTATTCACATGATGGTCAG 717  
Db 591 TGTATGCGCTTCTGCTGCGCTTTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 650  
Qy 718 GAGCTGTATCAAGCCAGAGGAGAACCTCATGAGGACAGGACAGCAGCCGAGCCAGCGTC 777  
Db 651 CCGCTGTGCGCGCAGGATGGCGGAGAGCTGTGCGCCAGGAGCGCGCTGGCAAGGC 710  
Qy 778 CATCGGAGCATCTACTGTGTGTGGCTCTTCCACCTCTGTTTGTGCGCTTCCATAT 837  
Db 711 GCGCGCATGCGCGTGGTGGTGGCTGCTCTTTGGCATCAGCTTCTGCTGCTTTTTCACAT 770  
Qy 838 CACTCGCTCTTCTACCTCACCATCTGCTTTCTTCTCAGGACTGCCAGCTCTTGAT 897  
Db 771 CACCAAGACACCTACCTGCGAGTGGCTCGACGCGCGGCTGCCCTGCATGTATTGGA 830  
Qy 898 GGCAGCCAGTGTGGCTTACAAGATATGAGGCGCTCTGTGTGAGTGTGAGCAGCTGCTCAA 957  
Db 831 GCGCTTTGACGCGCTTACAAGGACGCGCGCGCTTTGCCAGTGGCCAAACAGCGTCTGGA 890  
Qy 958 CCCAGTCTCTGACT 971  
Db 891 CCCATCCTCTTCT 904

RESULT 7  
US-08-513-974B-41  
; Sequence 41, Application US/08513974B  
; Patent No. 6114139  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Ohtaki, Tetsuya  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Ohgi, Kazuhiro  
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
; NUMBER OF SEQUENCES: 380  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,974B

```

SEQUENCE CHARACTERISTICS:
LENGTH: 984 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-513-974B-41

Query Match      13.8%; Score 148.8; DB 3; Length 984;
Best Local Similarity 50.8%; Pred. No. 1.6e-31;
Matches 434; Conservative 0; Mismatches 412; Indels 8; Gaps 3;

QY   126 CTGCCAGTCTTCAGAGAAGTACAAAGATCTACCTCTCCGTGGCCTACAGTATCATCTTT 185
      |||| | ||| | |||| | ||| | |||| | ||| | |||| | ||| | |||| |
Db    51 CTCGCTCTACCGTGAGGATTTCACGCAGACTGCTGAACCCCGGTATATCGGTGGTGCT 110

QY   186 TATCCTTAGGGCTGCCACTAATGGCACTGCTTGTGGCACTCTCTGGGCCCCAACCAAGCG 245
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db    111 GGTGGTGGGGCTGCCATGAACTATCGGTATCGCCAGATCTGGCGATCCC GCCGGAC 170

QY   246 CTGGAGCTGTGCCACCACCTATCTGTGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCT 305
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db    171 CCTGACCCGTTCCGCTGTGTACACCCTGAACCTGGCACTGGCGGACCTGATGTATGCCCTG 230?

QY   306 AT---TGCCCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCT 362
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db    231 TTCATACCCCTACTTATCTATACTACGCGCAGAGGGGACCACCTGGCCCTTCGGNAGACT 290

QY   363 GCCTCTGCAAGCTGGTGACCTTCCTGTTCTTATATCAACCTTTTACGGCGACGATCCTCTGCTGT 422
      |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

```

291	Db	CGCCTGCGCGCTTTGTACGCTTCTCTTCTATGCGCAATCTACATGCGACGATCCTGTTCCCT	350
423	QY	GACCTGCATCTCTGTGCACACAGTTCTCCTAGGTGTGTGCGCACCCACTGTGTTGCGCTGCCCTA	482
351	Db		410
483	QY	CC---GGAGCCGCGACGGCATGCGCTGGCGACACAGCACACCTGTGGGCGCTTGTGTGGTCTCT	539
411	Db		470
540	QY	CGAGTGTCTGCGCCACACATGGCGCTCTCCACACAGGACTACATCAATGCGCAGATGATCTG	599
471	Db		530
600	QY	GTATGACATGACCGACCGCAAGAGAAATTTTGATCGGCTTTTTGCTACGGCATAGTTCTGAC	659
531	Db		590
660	QY	ATTGTCTGGCT--TTCCTTTCCCTCTGGTGCTAATTTGGTGTGCTAATTCACGTATGGTGCAG	717
591	Db		650
718	QY	GAGCCTGATCAAGCGCAGAGGAGAACTCATGAGGACAGCACACAGCCGCGGACGATGTC	777
651	Db		710
778	QY	CATCGGGACATCTTACTGTGTGTGGCTCTTTCACCCCTCTGTTTGTGGCTTCCATAT	837
711	Db		770
838	QY	CAGTGCCTCTTCTACTTCACCATCTGCTTTCTGCTTCTTCAGAGCTGCGAGCTTCTGAT	897
771	Db		830
898	QY	GCAGCGCAGTGTGGCTACAAAGATATGAGGCGCTCTGGTGAGTGTGAGCAGCTGCCTCAA	957
831	Db		890
958	QY	CCAGCTCTGTACT	971
891	Db	CCCCATCTCTTCT	904

RESULT 8  
 US-08-513-974B-370  
 ; Sequence 370, Application US/08513974B  
 ; Patent No. 6114139  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hinuma, Shuji  
 ; APPLICANT: Hosoya, Masaki  
 ; APPLICANT: Fujii, Ryo  
 ; APPLICANT: Ohtaki, Tetsuya  
 ; APPLICANT: Fukusumi, Shoji  
 ; APPLICANT: Ohgi, Kazuhiro  
 ; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR PROTEIN,  
 ; TITLE OF INVENTION: PRODUCTION, AND USE THEREOF  
 ; NUMBER OF SEQUENCES: 380  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/513,974B  
 ; FILING DATE: 14-SEP-1995  
 ; CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP95/01599  
FILING DATE: 10-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-093989  
FILING DATE: 19-AUG-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-057186  
FILING DATE: 16-MAR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-007177  
FILING DATE: 20-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-326611  
FILING DATE: 28-DEC-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-270017  
FILING DATE: 02-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236357  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-236356  
FILING DATE: 30-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189274  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189273  
FILING DATE: 11-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-189272  
FILING DATE: 11-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S.  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 45753  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 370:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1020 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 28..1011  
US-08-513-974B-370

very Match 13.8%; Score 148.8; DB 3; Length 1020;  
st Local Similarity 50.8%; Pred. No. 1.7e-31;  
Matches 434; Conservative 0; Mismatches 412; Indels 8; Gaps 3;  
Qy 126 CTGCGAGTCTCAGAGGATACAGCAAGTCTACCTCTCCCTGCGCCACAGATATCATCTT 185  
Db 78 CTGGCTACCGTGAGGATTTCAAGCGACTGCTGCTAACCCCGGTATATCTCGGTGGTCT 137  
Qy 186 TATCCTAGGCTGCCATAATAGCACTGCTTTGGCACTCTTGGGCGCAACCAAGCG 245  
Db 138 GGTGGTGGGCTGCCATGACATCTCGCTCATTTGCCAGATCTGCGCATCCCGCGGAC 197  
Qy 246 CTGGAGTGTGCCACCACTATCTTGGTGAACCTGATGTGGCGACCTGCTTTATGTGCT 305  
Db 198 CCTGACCGTTCCGCTGTGTACACCTTGAACTGGCACTGCGGACCTGATGATGCTG 257  
Qy 306 AT---TGCCTTCTCATCATCTACTACTAGATACAGGTGGCCCTTTCGGGGAGCT 362  
Db 258 TTCACTACCCCTACTTATCTATACTACGCCAGAGGGGACCACTGGCCCTTTCGGAGACCT 317  
Qy 363 GCTCTGCAAGCTGTGGCACTTCTCTGTTCTATATCAACTTTACGGCAGCATCTCTGCTCT 422

Db 318 GCGCTGCGCGCTTTGTAGCCTTCTCTTCTATGCAATCTACATGCGCAGCATCTCTTCTCT 377  
Qy 423 GACCTGCATCTCTGTGCACCAAGTTCCCTAGGTGTGTGCGCACCACTGTGTGCTGCGCCCTA 482  
Db 378 CACCTGCATTTAGCTTCCAGCGCTACCTGGGCATCTGCGACCCCTGCTTCTGCTGCGACAA 437  
Qy 483 CC---GGACCCGCGAGGATGCTTGGCTGGGCACCAAGCACTGCGGCGCTTGGTGGTCT 539  
Db 438 GCGTGGAGTGGCGCTGCTGCTGGGTAGTGTGTGGAGTGTGTGGTGGTGTGTGACAGC 497  
Qy 540 CCAGCTGTGCCACACTGGCCTTCTCCACACAGGACTACATCAATGCGCAGAGATGCTG 599  
Db 498 CCAGTGTCTGCCACCGCAGTCTTTGCTGCCACAGGATCCAGCGCAACCCGACCTGTGTG 557  
Qy 600 GTATGACATGACCAAGCAAGAAATTTTGTGCTGCTTTTTCCTACGGCATAGTCTTGAC 659  
Db 558 CTAGCACTGAGCCCACTCCCTCTCTACTCGCTACTGCGCTATGCTATGCTATGCGCCTCAC 617  
Qy 660 ATTGCTGGCT--TTCTTTCCCTCCTTGGTCAATTTGGTGTGCTTATTCATGATGGTCA 717  
Db 618 GGTATCGGCTTCTTGTGCTGCTTCTATAGCCTTACTGGCTTGTATTGCTGCGATGGCCG 677  
Qy 718 GAGCTGATCAAGCCAGAGGAGAACCTCATGAGGACAGGCAACACAGCCCGAGCGAGTGC 777  
Db 678 CCGCTGTGTGCCAGGATGGCCCAAGGAGTCTGTGTGCCCAAGAGCGCGCAGCAAGGC 737  
Qy 778 CATCCGACCATCTACTTGTGTGGCTCTTACCCCTCTGTGTGTGCGCTTGTGCGCTTCCATAT 837  
Db 738 GGCTGATGCTGTGTGGTGGCGAGCTGCTTTGCCATCAGCTTCTGCGCTTCCCAT 797  
Qy 838 CACTGCTCTCTTCTACCTCACTCTGCTTCTCTTCTCAGGACTGCGCAGCTTGTGAT 897  
Db 798 CACCAAGACAGCTTCTTGTGCTGCGCTTCCACGCGGCTGCTCTTGGCCTGTGCTGA 857  
Qy 898 GGCAGCGAGTGTGCGCTACAGATATGAGGCGCTCTGCTGAGTGTGAGCAGCTGCTCAA 957  
Db 858 GACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGA 917  
Qy 958 CCCAGTCTCTTCT 971  
Db 918 CCCATCTCTTCT 931  
RESULT 9  
US-08-097-938-1  
; Sequence 1, Application US/08097938  
; Patent No. 5629174  
; GENERAL INFORMATION:  
; APPLICANT: SUNDELIN, JOHAN  
; APPLICANT: SCARBOROUGH, ROBERT M.  
; TITLE OF INVENTION: RECOMBINANT CL40 RECEPTOR AND ITS  
; TITLE OF INVENTION: AGONISTS AND ANTAGONISTS  
; NUMBER OF SEQUENCES: 59  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
; CITY: Washington, D.C.  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/097,938  
; FILING DATE: 26-JUL-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURASHIGE, KATE H.  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 22803-20006.00



## TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030

## INFORMATION FOR SEQ ID NO: 1:

## SEQUENCE CHARACTERISTICS:

LENGTH: 1475 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 232..1416

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 232

-08-097-938-1

Query Match 8.5%; Score 92; DB 1; Length 1475;

Best Local Similarity 56.5%; Pred. No. 7.8e-16;

Matches 192; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 134 TCTCAGAGAGTACACAGCAAGTCTCTCCCTGGCGCTACAGTATCATCTTTATCCCTAG 193

Db 431 TCACCGGAAGCTGACACCGCTTTCTCCGGTCTGTACATATTGTTTGTGATTG 490

QY 194 GCGTGCACCTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCT 253

Db 491 GTTTGCCAGTAGTATGCGCTCTGGATCTCTTTCGGAAGCAAGAAACACC 550

QY 254 GTGCCACCACTATCTGTTGGAACCTGATGGTGGCGGACCTGCTTTATGCTATTGCCCT 313

Db 551 CGCCGCTGATTTACATGGCAACCTGGCCCTTGGATCTCTTTCGGAAGCAAGAAACACC 610

QY 314 TCTCATCATCACTACTCACTAGATG---ACAGGTGGCCCTTTCGGGGAGCTCTGTGCA 370

Db 611 CCCTGAAGATCTCTACCACTACATGGCAACACTGGGTCTACGGGGAGCCCTGTGCA 670

QY 371 AGCTGGTGCACCTTCTCTATATCAACCTTTACGGCAGCATCTCTGCTGACCTGCA 430

Db 671 AGGTGCTCATGGCTTTTCTATGGAATGATGTTTCTCCATCTTTCATGACCTGCC 730

QY 431 TCCTCTGTGCACCACTTCTCTAGGTGTGGCCACCACTGTG 470

Db 731 TCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATGGG 770

## RESULT 10

US-08-476-000-1

Sequence 1, Application US/08476000

Patent No. 5716789

## GENERAL INFORMATION:

APPLICANT: SUNDELIN, JOHAN

APPLICANT: SCARBOROUGH, ROBERT M.

TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20006-1812

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,000

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/390,301

FILING DATE: 25-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: ADLER, REID G.

REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: 2803-0006.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

TELEX: 90-4030

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1475 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 232..1416

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 232

US-08-476-000-1

Query Match 8.5%; Score 92; DB 1; Length 1475;

Best Local Similarity 56.5%; Pred. No. 7.8e-16;

Matches 192; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

QY 134 TCTCAGAGAGTACACAGCAAGTCTCTCCCTGGCGCTACAGTATCATCTTTATCCCTAG 193

Db 431 TCACCGGAAGCTGACACCGCTTTCTCCGGTCTGTACATATTGTTTGTGATTG 490

QY 194 GCGTGCACCTAAATGGCACTGTCTTGTGGCACTCTCTGGGGCCAAACCAAGCGCTGGAGCT 253

Db 491 GTTTGCCAGTAGTATGCGCTCTGGATCTCTTTCGGAAGCAAGAAACACC 550

QY 254 GTGCCACCACTATCTGTTGGAACCTGATGGTGGCGGACCTGCTTTATGCTATTGCCCT 313

Db 551 CGCCGCTGATTTACATGGCAACCTGGCCCTTGGCGACCTCTCTCTGTCTATCTGGTTCC 610

QY 314 TCTCATCATCACTACTCACTAGATG---ACAGGTGGCCCTTTCGGGGAGCTGCTGTGCA 370

Db 611 CCCTGAAGATCTCTACCACTACATGGCAACACTGGGTCTACGGGGAGCCCTGTGCA 670

QY 371 AGCTGGTGCACCTTCTCTATATCAACCTTTACGGCAGCATCTCTGCTGACCTGCA 430

Db 671 AGGTGCTCATGGCTTTTCTATGGAATGATGTTTCTCCATCTTTCATGACCTGCC 730

QY 431 TCCTCTGTGCACCACTTCTCTAGGTGTGGCCACCACTGTG 470

Db 731 TCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATGGG 770

## RESULT 11

US-08-472-840-1

Sequence 1, Application US/08472840

Patent No. 5763575

## GENERAL INFORMATION:

APPLICANT: SUNDELIN, JOHAN

APPLICANT: SCARBOROUGH, ROBERT M.

TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS

TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,840

FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:

NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1475 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

NAME/KEY: CDS  
LOCATION: 232..1416  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 232

US-08-472-840-1

Query Match 8.5%; Score 92; DB 1; Length 1475;  
Best Local Similarity 56.5%; Pred. No. 7.8e-16;  
Matches 192; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

Qy 134 TCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATTCCTAG 193  
Db 431 TCACGGGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATTCCTAG 490  
Qy 194 GCCTGCCACTAAATGGCACTGTCTGTGGCACTCTCTGGCCCAACCAAGCGCTGGAGCT 253  
Db 491 GTTGGCCAGTAAATGGCACTGTCTGTGGCACTCTCTGGCCCAACCAAGCGCTGGAGCT 550  
Qy 254 GTGCCACCACTATCTGTGTGAACCTGTGTGGCCCAACCAAGCGCTGGAGCT 313  
Db 551 CGCGCGTATTTACATGCGCAACCTGTGTGGCCCAACCAAGCGCTGGAGCT 610  
Qy 314 TCCTCATCATCACCCTACTACTAGATG---ACAGTGGCCCTTCGGGGAGCTGCTGCA 370  
Db 611 CCCTGAAGATCTCTTACCACTACATGCGCAACCTGTGTGGCCCAACCAAGCGCTGGAGCT 670  
Qy 371 AGCTGGTGCACCTCTCTGTCTATATCAACCTTTACGGCAGCATCTCTGTGTGCTGCTGCA 430  
Db 671 AGGTGCTCATTTGCGCTTTTCTATGTAACATGTAATGCTCCATCTCTCTCATGACCTGCC 730  
Qy 431 TCTCTGTGCACAGTCTCTAGTGTGTGCCACCCACTGTG 470  
Db 731 TCAGGCGTCAGAGTACTGGGTGATCGTGAACCCCATGGG 770

RESULT 12  
US-08-476-976-1  
Sequence 1, Application US/08476976  
Patent No. 5874400  
GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,976  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1475 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
NAME/KEY: CDS  
LOCATION: 232..1416  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 232

US-08-476-976-1

Query Match 8.5%; Score 92; DB 2; Length 1475;  
Best Local Similarity 56.5%; Pred. No. 7.8e-16;  
Matches 192; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

Qy 134 TCTCAGAGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATTCCTAG 193  
Db 431 TCACGGGAAGTACAGCAAGTCTACCTCTCCCTGGCCCTACAGTATCATCTTTATTCCTAG 490  
Qy 194 GCCTGCCACTAAATGGCACTGTCTGTGGCACTCTCTGGCCCAACCAAGCGCTGGAGCT 253  
Db 491 GTTGGCCAGTAAATGGCACTGTCTGTGGCACTCTCTGGCCCAACCAAGCGCTGGAGCT 550  
Qy 254 GTGCCACCACTATCTGTGTGAACCTGTGTGGCCCAACCAAGCGCTGGAGCT 313  
Db 551 CGCGCGTATTTACATGCGCAACCTGTGTGGCCCAACCAAGCGCTGGAGCT 610  
Qy 314 TCCTCATCATCACCCTACTACTAGATG---ACAGTGGCCCTTCGGGGAGCTGCTGCA 370  
Db 611 CCCTGAAGATCTCTTACCACTACATGCGCAACCTGTGTGGCCCAACCAAGCGCTGGAGCT 670  
Qy 371 AGCTGGTGCACCTCTCTGTCTATATCAACCTTTACGGCAGCATCTCTGTGTGCTGCTGCA 430  
Db 671 AGGTGCTCATTTGCGCTTTTCTATGTAACATGTAATGCTCCATCTCTCTCATGACCTGCC 730  
Qy 431 TCTCTGTGCACAGTCTCTAGTGTGTGCCACCCACTGTG 470  
Db 731 TCAGGCGTCAGAGTACTGGGTGATCGTGAACCCCATGGG 770

RESULT 13  
US-08-474-410-1  
Sequence 1, Application US/08474410  
Patent No. 6043212  
GENERAL INFORMATION:

Query Match 8.5%; Score 92; DB 3; Length 1475;  
Best Local Similarity 56.5%; Pred. No. 7.8e-16;  
Matches 192; Conservative 0; Mismatches 145; Indels 3; Gaps 1;

Query Match

Query Match 8.5%; Score 92; DB 4; Length 1475;  
Best Local Similarity 56.5%; Pred. NO. 7.8e-16;  
Matches 192; Conservative 0; Mismatches 145 Indels

### Query Match

RESULT 15

STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,000  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 23-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2732 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 73..1269  
US-08-476-000-60

[illegible]

Search completed: May 26, 2003, 22:54:17  
Job time : 73 secs

GenCore version: 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 21:24:13 ; Search time 199 Seconds  
(without alignments)  
7179.599 Million cell updates/sec

Title: US-10-023-586b-1  
Perfect score: 1082  
Sequence: 1 atgtctgtccatttgcctcc.....ccagggttgacagatctgg 1082

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues  
Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1006	93.0	1017	9	US-09-885-453-3
2	510	47.1	510	9	US-09-782-974C-13
3	222	20.5	551	9	US-09-791-932-32
c 4	178.6	16.5	1429	9	US-09-077-173A-1
5	97.8	9.0	1543	9	US-09-891-138A-1
6	97.4	9.0	1773	9	US-10-112-599A-3
7	95.6	8.8	1892	9	US-09-900-699A-1
8	95.2	8.8	422	9	US-09-779-679-23
9	89	8.2	1526	10	US-09-728-952-40
10	88.4	8.2	1020	10	US-09-788-133-1
11	87.4	8.1	1679	9	US-09-104-063-5
12	86	7.9	1313	10	US-09-728-422-1
c 13	85.8	7.9	348	10	US-09-884-761-21546
14	83.6	7.7	1829	10	US-09-823-114-7
15	83.6	7.7	2218	10	US-09-214-904-3
16	83.6	7.7	2219	10	US-10-112-599A-1
17	83.2	7.7	1164	9	US-10-228-264-3
18	83.2	7.7	1365	7	US-08-899-112-29
19	82.6	7.6	2693	9	US-09-850-948-1

20	82	7.6	1805	10	US-09-823-114-18	Sequence 18, Appli
21	82	7.6	1829	9	US-09-905-186A-9	Sequence 9, Appli
22	82	7.6	1829	9	US-09-905-186A-10	Sequence 10, Appli
23	82	7.6	1829	9	US-09-905-186A-11	Sequence 11, Appli
24	82	7.6	2534	9	US-10-087-345A-22	Sequence 22, Appli
25	81.8	7.6	2602	9	US-09-905-186A-1	Sequence 1, Appli
26	80.8	7.5	1102	9	US-09-870-759-143	Sequence 143, App
27	80.8	7.5	1670	10	US-09-880-107-2143	Sequence 2143, Ap
28	80.8	7.5	1737	9	US-09-104-063-3	Sequence 3, Appli
29	80.4	7.4	1829	9	US-09-905-186A-7	Sequence 7, Appli
30	80.4	7.4	1829	9	US-09-905-186A-8	Sequence 8, Appli
31	79.6	7.4	2427	10	US-09-254-783A-2	Sequence 2, Appli
32	79.6	7.4	2427	12	US-10-152-058-2	Sequence 2, Appli
c 33	78.4	7.2	442	10	US-09-864-761-4808	Sequence 4808, Ap
34	77.6	7.2	1107	9	US-10-212-980-3	Sequence 3, Appli
35	77.6	7.2	1107	9	US-10-228-264-1	Sequence 1, Appli
36	77.6	7.2	1417	9	US-10-007-132-3	Sequence 3, Appli
37	77.6	7.2	1417	9	US-09-771-287-3	Sequence 3, Appli
38	77.4	7.2	1586	10	US-09-104-792-1	Sequence 1, Appli
39	77	7.1	1098	9	US-09-850-948-5	Sequence 5, Appli
40	77	7.1	1697	9	US-10-109-533A-1	Sequence 1, Appli
41	76.4	7.1	1404	10	US-09-805-628-3	Sequence 3, Appli
42	76.4	7.1	1419	10	US-09-805-628-1	Sequence 1, Appli
43	76.2	7.0	318	10	US-09-728-445-777	Sequence 777, App
44	76.2	7.0	1014	9	US-10-023-775B-1	Sequence 1, Appli
45	76.2	7.0	1014	9	US-10-270-144-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-885-453-3  
; Sequence 3, Application US/09885453  
; Publication No. US20030088080A1  
; GENERAL INFORMATION:  
; APPLICANT: Communi, Didier  
; TITLE OF INVENTION: RECEPTOR GPCRxl0  
; FILE REFERENCE: 9409/2082  
; CURRENT APPLICATION NUMBER: US/09/885,453  
; CURRENT FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: US 09/885,453  
; PRIOR FILING DATE: 2001-06-21  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1017  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: DNA sequence  
; LOCATION: (1)..(1017)  
; OTHER INFORMATION: GPCRx6 DNA sequence  
US-09-885-453-3

Query Match	93.0%;	Score	1006;	DB	9;	Length	1017;
Best Local Similarity	99.9%;	Pred. No.	1.4e-306;				
Matches	1017;	Conservative	0;	Mismatches	0;	Indels	1;
Gaps	1;						
QY	1	ATGCTGTCCATTGTTGTTCTTCCAGGGGAGCAGAGCGGAGCCGTCGTGGAGTCTG	60				
Db	1	ATGCTGTCCATTGTTGTTCTTCCAGGGGAGCAGAGCGGAGCCGTCGTGGAGTCTG	60				
QY	61	CTCCTTGAGGAGCCCTCCCGGACATGGAGAGTGGAGATGAATACATCACAGGAACAA	120				
Db	61	CTCCTTGAGGAGCCCTCCCGGACATGGAGAGTGGAGATGAATACATCACAGGAACAA	120				
QY	121	GGTCTCTGCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC	180				
Db	121	GGTCTCTGCAGTTCTCAGAGAGTACAGCAAGTCTACCTCTCCCTGGCCTACAGTATC	180				
QY	181	ATCTTTATCTAGGCGTGCACCTAAATGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC	240				
Db	181	ATCTTTATCTAGGCGTGCACCTAAATGCACTGTCTTGTGGCACTCTCTGGGGCCAAACC	240				

Db 181 ATCTTTATCTAGGCTGCCACTAAATGGCACTCTCTGTGGCACTCTCTGGGGCAAAACC 240  
Qy 241 AAGCGTGGAGCTGTGCCACCACTATCTGTGAACCTGTATGGTGGCGACCTGCTTTAT 300  
Db 241 AAGCGTGGAGCTGTGCCACCACTATCTGTGAACCTGTATGGTGGCGACCTGCTTTAT 300  
Qy 301 GTGCTATTGCCCTTCTCATCATCACTACCTACTAGATGACAGGTGGCCCTTCCGGGAG 360  
Db 301 GTGCTATTGCCCTTCTCATCATCACTACCTACTAGATGACAGGTGGCCCTTCCGGGAG 360  
Qy 361 CTGCTCTGCAAGCTGGTGCACCTCTCTGTCTATATCAACCTTTAGCGCAGCATCTGCTG 420  
Db 361 CTGCTCTGCAAGCTGGTGCACCTCTCTGTCTATATCAACCTTTAGCGCAGCATCTGCTG 420  
Qy 421 CTGACCTGCATCTCTGTGCACCACTCTCTAGTGTGTGCCACCACTGTGTGCTGCC 480  
Db 421 CTGACCTGCATCTCTGTGCACCACTCTCTAGTGTGTGCCACCACTGTGTGCTGCC 480  
Qy 481 TACCGGACCCGAGCATGCTGTGGTGGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 540  
Db 481 TACCGGACCCGAGCATGCTGTGGTGGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 540  
Qy 541 CAGCTGTGCCACCACTGGCTCTCCACACAGGACTACATCAATGCCAGCATCTGCTG 600  
Db 541 CAGCTGTGCCACCACTGGCTCTCCACACAGGACTACATCAATGCCAGCATCTGCTG 600  
Qy 601 TATGACATGACAGCAGCAAGAGATTTTGTATCGGCTTTTGTGCTACGGCATAGTCTGACA 660  
Db 601 TATGACATGACAGCAGCAAGAGATTTTGTATCGGCTTTTGTGCTACGGCATAGTCTGACA 660  
Qy 661 TTGCTGTGCTTTCTTCCCTCCTTGGTCAATTTTGGTGTGCTATTTCATGATGGTCAGGAG 720  
Db 661 TTGCTGTGCTTTCTTCCCTCCTTGGTCAATTTTGGTGTGCTATTTCATGATGGTCAGGAG 720  
Qy 721 CCTGATCAAGCAGCAGGAGACCTCATGAGCAGCAGGCAACACAGCAGCAGCAGCAGCAG 780  
Db 721 CCTGATCAAGCAGCAGGAGACCTCATGAGCAGCAGGCAACACAGCAGCAGCAGCAGCAG 780  
Qy 781 CCGGACCATCTACTGTGTGTGGCTCTTCCACCTCTGTGTTTGTGGCTTCCATATAC 840  
Db 781 CCGGACCATCTACTGTGTGTGGCTCTTCCACCTCTGTGTTTGTGGCTTCCATATAC 840  
Qy 841 TCGCTCTCTTACCTCACCATCTGCTTCTGCTTCTTCCAGCAGTGGCAGCTTGTATGGC 900  
Db 841 TCGCTCTCTTACCTCACCATCTGCTTCTGCTTCTTCCAGCAGTGGCAGCTTGTATGGC 900  
Qy 901 AGCCAGTGTGGCTTACAAGATATGAGGCTCTGTGAGTGTGAGCAGCTGCTTCAACCC 960  
Db 901 AGCCAGTGTGGCTTACAAGATATGAGGCTCTGTGAGTGTGAGCAGCTGCTTCAACCC 960  
Qy 961 AGTCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAGGCTCTTCCAGAACTGA 1018  
Db 961 AGTCTGTACTTTCTTTCAAGGGGGGCAAAATAGAGTCAGGCTCTTCCAGAACTGA 1018

## RESULT 2

US-09-782-974C-13  
; Sequence 13, Application US/09782974C  
; Publication No. US20030082534A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogel, Gabriel  
; APPLICANT: Lind, Peter  
; APPLICANT: Wood, Linda S.  
; APPLICANT: Parodi, Luis A.  
; TITLE OF INVENTION: No. US20030082534A1e1 G Protein Coupled Receptor  
; FILE REFERENCE: 411USPHRM311  
; CURRENT APPLICATION NUMBER: US/09/782,974C  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: 60/165,838  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR APPLICATION NUMBER: 09/714,449  
; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: 60/198,568

; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 60/166,071  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: 60/166,678  
; PRIOR FILING DATE: 1999-11-19  
; PRIOR APPLICATION NUMBER: 60/173,396  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/184,129  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: 60/185,421  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/185,554  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/186,530  
; PRIOR FILING DATE: 2000-03-02  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 13  
; LENGTH: 510  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-782-974C-13

Query Match 47.1%; Score 510; DB 9; Length 510;  
Best Local Similarity 100.0%; Pred. No. 2e-150;  
Matches 510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 247 TGGAGCTGTGCCACCACTATCTGTGAACCTGTATGGTGGCGACCTGCTTTATGCTGA 306  
Db 1 TGGAGCTGTGCCACCACTATCTGTGAACCTGTATGGTGGCGACCTGCTTTATGCTGA 60  
Qy 307 TTGCGCTTCTCATCATCACTACCTACTAGATGACAGGTGGCCCTTCCGGGAGCTGCTC 366  
Db 61 TTGCGCTTCTCATCATCACTACTACTAGATGACAGGTGGCCCTTCCGGGAGCTGCTC 120  
Qy 367 TGAAGCTGTGGTGCATCTCTGCTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACC 426  
Db 121 TGAAGCTGTGGTGCATCTCTGCTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACC 180  
Qy 427 TGCATCTGTGGCACCAGTTCTCTAGTGTGTGCCACCACTGTCTGCTGCCCTTACCGG 486  
Db 181 TGCATCTGTGGCACCAGTTCTCTAGTGTGTGCCACCACTGTCTGCTGCCCTTACCGG 240  
Qy 487 ACCGCGAGGATGCTGCTGGTGGCACCAGCACCACCTGGGCGCTGGTGGTCTCCAGCTG 546  
Db 241 ACCGCGAGGATGCTGCTGGTGGCACCAGCACCACCTGGGCGCTGGTGGTCTCCAGCTG 300  
Qy 547 CTGCGCCACACTGGCTTCTCCACACAGGACTACATCAATGCCAGATGATCTGCTATGAC 606  
Db 301 CTGCGCCACACTGGCTTCTCCACACAGGACTACATCAATGCCAGATGATCTGCTATGAC 360  
Qy 607 ATGACAGGCAAGAGAATTTTGATCGGCTTTTGCCTACGGCATAGTTCTGACATTTGCT 656  
Db 361 ATGACAGGCAAGAGAATTTTGATCGGCTTTTGCCTACGGCATAGTTCTGACATTTGCT 420  
Qy 667 GGCCTTCTTCCCTTCCCTGCTGCTATTTGGTGTGCTATTTACTGATGGTCAGGAGCCTGAT 726  
Db 421 GGCCTTCTTCCCTTCCCTGCTGCTATTTGGTGTGCTATTTACTGATGGTCAGGAGCCTGAT 480

## RESULT 3

US-09-791-932-32/c  
; Sequence 32, Application US/09791932  
; Publication No. US20030003451A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogel, Gabriel  
; APPLICANT: Parodi, Luis A.  
; APPLICANT: Hiebsch, Ronald R.

APPLICANT: Lind, Peter  
APPLICANT: Kaytes, Paul S.  
APPLICANT: Ruff, Valerie  
APPLICANT: Huff, Rita M.  
APPLICANT: Wood, Linda S.  
TITLE OF INVENTION: No. US2003003451A1el G Protein-Coupled Receptors Cross-Referenced  
FILE REFERENCE: 00325.US1  
CURRENT APPLICATION NUMBER: US/09/791,932  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: 60/184,305  
PRIOR FILING DATE: 2000-02-23  
PRIOR APPLICATION NUMBER: 60/184,304  
PRIOR FILING DATE: 2000-02-23  
PRIOR APPLICATION NUMBER: 60/184,303  
PRIOR FILING DATE: 2000-02-23  
PRIOR APPLICATION NUMBER: 60/184,397  
PRIOR FILING DATE: 2000-02-23  
PRIOR APPLICATION NUMBER: 60/184,247  
PRIOR FILING DATE: 2000-02-23  
PRIOR APPLICATION NUMBER: 60/188,880  
PRIOR FILING DATE: 2000-03-13  
PRIOR APPLICATION NUMBER: 60/217,369  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/217,370  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/218,492  
PRIOR FILING DATE: 2000-07-20  
PRIOR APPLICATION NUMBER: 60/186,810  
PRIOR FILING DATE: 2000-03-03  
PRIOR APPLICATION NUMBER: 60/188,064  
PRIOR FILING DATE: 2000-03-09  
PRIOR APPLICATION NUMBER: 60/186,457  
PRIOR FILING DATE: 2000-03-02  
PRIOR APPLICATION NUMBER: 60/213,861  
PRIOR FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: 60/194,344  
PRIOR FILING DATE: 2000-04-03  
PRIOR APPLICATION NUMBER: 60/218,337  
PRIOR FILING DATE: 2000-07-14  
NUMBER OF SEQ ID NOS: 184  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 32  
LENGTH: 551  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-791-932-32

Query Match 20.5%; Score 222; DB 9; Length 551;  
Best Local Similarity 100.0%; Pred. No. 1.2e-59;  
Matches 222; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGTCCATTTGCTTCCAGGGAAGCAGAGCGGAGCGCTGCTGAGCTCTG 60  
DB 222 ATGCTGTCCATTTGCTTCCAGGGAAGCAGAGCGGAGCGCTGCTGAGCTCTG 163

QY 61 CTCCTGGAGGAGCCCTCCCGGACATGAGAGAGGTGGACATGAATACATACAGGAACAA 120  
DB 162 CTCCTGGAGGAGCCCTCCCGGACATGAGAGAGGTGGACATGAATACATACAGGAACAA 103

QY 121 GGTCTGTCCAGTTCTCAGAGAGTACAGCAAGTCTACTCTCCCTGGCCCTACAGTATC 180  
DB 102 GGTCTGTCCAGTTCTCAGAGAGTACAGCAAGTCTACTCTCCCTGGCCCTACAGTATC 43

QY 181 ATCTTTATCTAGGCTGCCACTTAATGGCACTGTCTTCTGG 222  
DB 42 ATCTTTATCTAGGCTGCCACTTAATGGCACTGTCTTCTGG 1

RESULT 4

US-09-077-173A-1  
Sequence 1, Application US/09077173A  
Publication No. US20030082674A1  
GENERAL INFORMATION:

APPLICANT:  
TITLE OF INVENTION: RECEPTOR AND NUCLEIC ACID MOLECULE ENCODING  
NUMBER OF SEQUENCES: 4  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/077,173A  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/BE 96/00123  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1429 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 181..1275  
US-09-077-173A-1

Query Match 16.5%; Score 178.6; DB 9; Length 1429;  
Best Local Similarity 54.0%; Pred. No. 8.6e-46;  
Matches 481; Conservative 0; Mismatches 394; Indels 16; Gaps 5;

QY 126 CTGCCAGTTCTCAGAGAAGTCAACGAAGTCACTCTCCCTGGCTACAGTATCATCTT 185  
DB 258 CTGTTGGTTTGATGAGGATTTCAAGTTTCATCTGCTGCCTGTGAGCATGCGATTT 317

QY 186 TATCTAGGCTGCCACATAATGGCACTCTCTTTGGGCACTCTCTGGGGCAACACCAAGG 245  
DB 318 TGTGTGGCTTTGGGCTTTAAGCCCCCAACCTATGGCTCTTCATCTCCGCTCCGACC 377

QY 246 CTGGAGCTGTGCCACCACCTATCTGTGTGAACCTGATGGTGGCGGACCTGCTTTATCTGCT 305  
DB 378 CTGGGATGCAAGCGCCACCTACATGTTCCACCTGGGATTTGTCAGACACCTTTGATGCT 437

QY 306 AT---TGCCCTTCTCATCATCATCACTACTACTAGATGACAGGTGGCCCTTCGGGGAGCT 362  
DB 438 GTCGCTGCCACCTCATCTACTATTTATGAGCCCAACCACTGSCCTTTGGCACTGA 497

QY 363 GCTCTGCAAGCTGTGCACTTCTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCT 422  
DB 498 GATCTGCAAGTTGCTGCGCTTTCTTTCTATTGGAACCTCTACTGCACTGCTCTTTTCT 557

QY 423 GACCTGCATCTGTGSCACAGTTCTCTAGGTGTGTGCCACCACTGTGTTCCGTGCCCTA 482  
DB 558 GACCTGCATACGGTGCACTGCTTCTGCTGCGGAGTTTGTGTTGCTGCTAGCCGCTG 617

QY 483 CCGGACCCCGCAGGATGCTGCTGGGCAACCACTGCTGGCCCTGCTGGTCTCTCCA 542  
DB 618 GGGCGCCCTCGCTCGCAGGCTTCTCTGCTGCGGAGTTTGTGTTGCTGCTAGCCGCTG 677

QY 543 GCTGCTGCCCACTGCGCTTCTCCCAACAGGACTACATCAATGAGGCAAGATGATGGA 602  
DB 678 CCTGCTGCCCACTGCTTCTTGTGTACAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAG 737

QY 603 TGACATGACCAAGCAAGATTTGATCGGCTTTTGGCTAGGCAATGTTCTGACATT 662  
DB 738 TGACACCACTCGGCTGGAAGAGTTTGACCACTATGTCACCTCGGCGGCTATGGG 797

QY 663 GTCTGGCTTT---CTTTCCCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 720  
DB 798 GCTGCTCTTTGGGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857

QY 721 CCTGATCAAGCCAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780  
DB 721 CCTGATCAAGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780





Db 500 CAACCTGGCCCTTAGCCGATGGCTGGCCACGACGACGCTGCTTCCAGAGTGCCAAAGTA 559  
Qy 333 ACTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGACCTTCTGTGTTCTA 392  
Db 560 CCGTATGAGAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGACCTTCTGTGTTCTA 619  
Qy 393 TATCAACCTTTAGCGAGCAGCTGCTGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452  
Db 620 CTACAATATGTTACCAAGCATCTTCACGCTCACCATGATGAGTGTTCACCGCTACATCGC 679  
Qy 453 TGTGTGCCACCCACTGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512  
Db 680 TGTGTGCCACCCCTGTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 739  
Qy 513 CAGCACCACCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 563  
Db 740 CATCTGTATCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 790

## RESULT 7

US-09-900-699A-1

; Sequence 1, Application US/09900699A

; Patent No. US20020162132A1

; GENERAL INFORMATION:

; APPLICANT: Brennan, Thomas J.

; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING DEZ ORPHAN

; FILE REFERENCE: R-173

; CURRENT APPLICATION NUMBER: US/09/900,699A

; PRIOR FILING DATE: 2002-05-22

; PRIOR APPLICATION NUMBER: US 60/262,137

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/251,815

; PRIOR FILING DATE: 2000-12-06

; PRIOR APPLICATION NUMBER: US 60/219,403

; PRIOR FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: US 60/216,253

; PRIOR FILING DATE: 2000-07-06

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 1892

; TYPE: DNA

; ORGANISM: Mus musculus

; US-09-900-699A-1

Query Match 8.8%; Score 95.6; DB 9; Length 1892;  
Best Local Similarity 50.4%; Pred. No. 1.4e-19;  
Matches 233; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

Qy 154 GTCTACCTCTCCCTGGCTACAGTATCATCTTTATCTAGGCTGCCACTAAATGGCACT 213  
Db 379 GTCTCTCTGGTGTATCTACAGTGTGGTGTCTTCCTCGCTCTCCCTAGGCAAGGGCTG 438  
Qy 214 GTCTTGTGGCACTCCTGGGGCAACCAAGCGCTGGAGCTGTGCCACCACTATCTGGT 273  
Db 439 GTGATTGTCTATCGCCACTTCAAGATGAAGAAGCGCTGAACACTGTGTGTGTCAAC 498  
Qy 274 AACCTGATGTGGCGGACCTGCTTTATGTGCTATGCTGCTTCTTCATCATCACTACTCA 333  
Db 499 CTGCTGTGGCGGACTTCTCTGTTCACACTCTTTTGGCGATGTCACATCACTACGCGGCC 558  
Qy 334 CTAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACCTCTCTGTTCTAT 393  
Db 559 ATGACTACCACTGGGTGTGGGAAGCCATGTGCAAGATCAACAACCTTCTTCTCTCAGC 618  
Qy 394 ATCAACCTTTACGGCAGCATCTCTGCTGACCTGCACTCTCTGTGCAACCACTTCTTAGGT 453  
Db 619 CACAACATGTACACAGGCTTCTCTGCTGACTGTCTATCATAGCTTGTACCGCTGATCTCC 678  
Qy 454 GTGTGCCACCACTGTGTGCTGCTGCTTACCGGACCGCGAGGATGCTGCTGGCGACC 513

Db 679 GTGCTGCTCCCGCTCTGGTCCAGAACCCGCGAGCATCGCGCTGGCTACATGACCTGC 738  
Qy 514 AGCACCACTGGGCCCTGGTGGTCTCCAGCTGCTGCCACACCTGGCTTCTCCCAACAG 573  
Db 739 TCGGCCCTCTGGGCTGGCTTCTTCTGAGCTCCCGCTGCTGCTTCTCCGGGACACC 798  
Qy 574 GACTACATCAATGCGCCAGATGATCTGGTATGACATGACACGAC 615  
Db 799 GCCAACATTTCATGGAAGATAAATCTGCTTCAACAACCTTCAGC 840

## RESULT 8

US-09-779-679-23

; Sequence 23, Application US/09779679

; Publication No. US20030082757A1

; GENERAL INFORMATION:

; APPLICANT: Taupier, Raymond J

; APPLICANT: Burgess, Catherine E

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Tchernev, Velizar T

; APPLICANT: Mishra, Vishnu S

; APPLICANT: Casman, Stacie

; APPLICANT: Ballinger, Robert

; APPLICANT: Vernet, Corine A

; APPLICANT: Li, Li

; APPLICANT: Spytek, Kimberly A

; APPLICANT: Andrew, David P

; APPLICANT: Mezes, Peter S

; TITLE OF INVENTION: No. US20030082757A1e1 Proteins and Nucleic Acids Encoding the

; FILE REFERENCE: 15966-661

; CURRENT APPLICATION NUMBER: US/09/779,679

; PRIOR FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: USSN 60/181045

; PRIOR FILING DATE: 2000-02-08

; PRIOR APPLICATION NUMBER: USSN 60/183191

; PRIOR FILING DATE: 2000-02-17

; PRIOR APPLICATION NUMBER: USSN 60/180,929

; PRIOR FILING DATE: 2000-02-08

; PRIOR APPLICATION NUMBER: USSN 60/219758

; PRIOR FILING DATE: 2000-07-20

; PRIOR APPLICATION NUMBER: USSN 60/181339

; PRIOR FILING DATE: 2000-02-09

; PRIOR APPLICATION NUMBER: USSN 60/181344

; PRIOR FILING DATE: 2000-02-09

; PRIOR APPLICATION NUMBER: USSN 60/221341

; PRIOR FILING DATE: 2000-07-26

; PRIOR APPLICATION NUMBER: USSN 60/181392

; PRIOR FILING DATE: 2000-02-09

; PRIOR APPLICATION NUMBER: USSN 60/219585

; PRIOR FILING DATE: 2000-07-20

; PRIOR APPLICATION NUMBER: USSN 60/181157

; NUMBER OF SEQ ID NOS: 65

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 23

; LENGTH: 422

; TYPE: DNA

; ORGANISM: Gallus gallus

; US-09-779-679-23

Query Match 8.8%; Score 95.2; DB 9; Length 422;  
Best Local Similarity 55.0%; Pred. No. 9.4e-20;  
Matches 187; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

Qy 169 GCCTACAGTATCATCTTTATCTAGGCTGCCACTAAATGGCACTGCTTGTGGCACTCC 228  
Db 52 GTCTTACAGTATGTTTGTCTCGGCTCATAGCCAACTGCTGTATCTACATTTT 111  
Qy 229 TGGGGCCAAACCAAGCGCTGGAGCTGCGCACCACTATCTGGTGAACCTGATGGTGCC 288  
Db 112 ACTTTTACATTGAAGTGGGAGACGAGACGACGATACATGCTGAATTTGGCGATATCG 171  
Qy 289 GACCTGCTTTATGTGCTATTGCGCTTCTCTCATCATCACTACTACTAGATGACAGGTGG 348

Db 172 GACCTGCTGTTTGTGTTTACGTTGCGCTTCAGGATCTATTACTTGGTGAGGAACG 231  
Qy 349 CCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTCTCTTATATCAACCTTTACGGC 408  
Db 232 CCCTTCGGAGAGCTGCTGTGCAAGATCTCCGTCAGCTGTTCTTACACCAATGATGACGG 291  
Qy 409 AGCATCCTGCTGCTGACCTGCACTCTCTGTGCAACAGTTCTCTAGGTGTGTGCCACCCACTG 468  
Db 292 AGCATTTCTATTCTGACCTGATCATGAGCTGGATGCTTCTCTGCGCATAGTGACACCTTT 351  
Qy 469 TGTTCGCTGCGCTTACGGACCCGAGGATGCTGCTGGTGG 508  
Db 352 CGCTCTAAGACTCTTTCGACCAAAAGGAACGACGAGGATCG 391  
  
RESULT 9  
US-09-728-952-40  
; Sequence 40, Application US/09728952  
; Patent No. US20020111302A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Goodrich, Ryle  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Yamazaki, Vicki  
; APPLICANT: Ujwal, Manusha L.  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. US20020111302A1el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/728,952  
; CURRENT FILING DATE: 2000-11-30  
; NUMBER OF SEQ ID NOS: 101  
; SOFTWARE: pc\_flgens Version 2.0  
; SEQ ID NO 40  
; LENGTH: 1526  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (684)..(1016)  
US-09-728-952-40  
  
Query Match 8.2%; Score 89; DB 10; Length 1526;  
Best Local Similarity 50.1%; Pred. No. 1.5e-17;  
Matches 302; Conservative 0; Mismatches 295; Indels 6; Gaps 3;  
  
Qy 158 ACCTCTCCCTGCGCTACAGTATCATCTTTATCTAGGCTGCCACTAAATGGCACTGTCT 217  
419 ACGGCTCTCTGTTCTACACAGCCCTCTGTTCTCAGTGCCTTGGAAACATCTTTGCC 478  
Qy 218 TGTGGCACTCTGGGGGCAAAACCAAGCGTGGAGCTGGCCACCACTATCTGTTGAGCC 277  
Db 479 TTTGCTTACCTGTCAAAAGAGAGAGAGATCAACTGCACAGGCATCTACCTGTGCACC 538  
Qy 278 TCATGGTGGCCGACCTGCTTTAT---GTGCTATTGCCCTTCTCATCATCACTACTCAC 334  
Db 539 TGGCTGTGCTGACCTGTCTTACCGTGGCCCTTACCGGGAAGGGTGGTGTATTATGTC 598  
Qy 335 TAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGGTGCACTTCTCTTCTATA 394  
Db 599 TGGGCTCCAGCTGGCTTTTCGCAAGGGCTCTGAGGCTGACGGCTTGTGCTCTACA 658  
Qy 395 TCAACCTTTACGGAGCATCTCTGCTGACCTCATCTCTGTGCAACAGTTCTCTAGGTG 454  
Db 659 CCGACACCTACGG-GGGGTCTCTCATGGCTGTGTGAGCGTGGACCATTTACCCAGCTG 717  
Qy 455 TGTGGCACCACTGTGTTGCTGCTTACCGGACCCGAGGATGCTGCTGGCTGGGCACCA 514

Db 718 TGGTCTGTGCCCACTGGGGCCCGTGCCTCCGACAGGCTGGCGCCCGCCAGGCTGTCTGG 777  
Qy 515 GCACCACCTGGCCCTGCTGGTGTCTCCAGCTGCTCCACACACTGCGCTTCTCCACACGG 574  
Db 778 TGGCATCTGGACCTTGGTGTCTGAGACGATGCCCTTCTCTTGTATGCCCATGACCA 837  
Qy 575 ACTACATCAATGGCCAGATGATCTGATATGACATGACGAGCAAGAAATTTTGTATCGGC 634  
Db 838 AGCCGCTGGTGGCAAGCTGGCTGCATGGAGTACAGCAGCATGGAGTCAGTCTCGGC 897  
Qy 635 TTTTTCCTACGGCATAGTTCTGACATTTGCTGGGCTTTCTTCCCTCCTTGG--TCATTT 692  
Db 898 TGCCCTCATGCTGCTGCTGGTGGCTTTGGCAATTTGGCTTCTGTGGGCCAGTGGGGATC 957  
Qy 693 TGGTGTGCTATTCACTGATGTCAGAGCCTGATCAAGCCAGAGAGAACTCATATGAGA 752  
Db 958 TGTCTCTATATGAAGATCACCTGGAAGCTGTGAGCAGACAGCTGGGAGAACCACTGAC 1017  
Qy 753 CAG 755  
Db 1018 CAG 1020  
  
RESULT 10  
US-09-788-133-1  
; Sequence 1, Application US/09788133  
; Patent No. US20020052001A1  
; GENERAL INFORMATION:  
; APPLICANT: GLAXO GROUP LTD  
; TITLE OF INVENTION: ASSAY  
; FILE REFERENCE: P79011  
; CURRENT APPLICATION NUMBER: US/09/788,133  
; CURRENT FILING DATE: 2001-02-20  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1020  
; TYPE: DNA  
; ORGANISM: homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1020)  
US-09-788-133-1  
  
Query Match 8.2%; Score 88.4; DB 10; Length 1020;  
Best Local Similarity 51.6%; Pred. No. 2e-17;  
Matches 228; Conservative 0; Mismatches 211; Indels 3; Gaps 1;  
  
Qy 161 TCTCCTGCGCTACAGTATCATCTTTATCTAGGCTGCCACTAAATGGCACTGTCTGT 220  
Db 101 TCGCCTCCTTCTACCTTCTGGATTTTATCTCTGGCTTTAGTTGGCAATACCTTGGCTCTGT 160  
Qy 221 GGCATCTCTGGGGCAAAACCAAGCGCTGGAGCTGTGCCACCACTATCTGTGGAACCTGA 280  
Db 161 GCTTTTTCATCCGAGACCAAGTCCGGGACCCCGCCCAACGTGTCTCTGATGATCTGG 220  
Qy 281 TGGTGGCGCACTGCTTTATGTGTAT---TGCCCTTCTCATCATCACTACTACTACCTAG 337  
Db 221 CCGTGGCGCACTTGTGCTGGCTGTGTCTGCCACCCGCTGTCTACCACTTCTCTGT 280  
Qy 338 ATGACAGTGGCCCTTCGGGAGCTGTCTGCAAGCTGGTGCACCTTCTGTCTATATCA 397  
Db 281 GGAACCACTGGCCATTTGGGAAATCGCATCCGCTTCTACCGGCTTCTCTTCTACCTCA 340  
Qy 398 ACCTTTACGGAGCATCTGCTGTGCTGACCTCTGTGACCACTGTCTAGTGTGT 457  
Db 341 ACATGACCGCAGCATCTACTTCTCCTACCTGATCAGCGCCGACCGTTTCTTGGCCATTG 400  
Qy 458 GCCACCCCACTGTGTGCTGCCCTTACCGGACCCGAGGATGCTGCTGGTGGGCAACCA 517  
Db 401 TGCACCCCGTCAAGTCCCTCAAGCTCCGAGGCCCTCTACGACACACCTGCTGCTGCT 460  
Qy 518 CCACCTGGGCCCTGTGTGCTCTCCAGCTGTGCCACACTGGCCTTCTCCACACAGGACT 577



Db 759 TATGCTGAACCTGGCGCTGCACAGATCTGCTATCTGACAGCCTCCCTTCTCCTGATTCA 818  
Qy 324 CACCTACTACTAGATGAGCTGCGCTTCGGGAGCTGCTGCGAGCTGGTGCACCTT 383  
Db 819 CTATATGCCAGTGGGAACTGATCTTTGGAGATTTCAATGTTATCCGCTT 878  
Qy 384 CCTGTTCTATATCAACCTTTACGGCAGCATCTGCTGCTGACCTGCATCTCTGTGCACCA 443  
Db 879 CAGCTTCCATTTCAACCTGTATAGCAGCATCTCTTCTCCTACCTGTTTTCAGCATCTTCCG 938  
Qy 444 GTTCTAGGTGTGCGCACCACTGTGTCTGCTGCGCTTACCGGACCCGAGCATGCGTG 503  
Db 939 CTACTGTGTGATCATTTACCAATGAGTGTCTTTTCCATTTCACAAACTCGATGTGCACT 998  
Qy 504 GCTGGGACACGACCACTGGCGCTGGTGTCTCTCCAGCTGCTGCCACACTGGCCTT 563  
Db 999 TGTAGCTGTCTGCTGGTGTGGATCATTTCACTGGTAGCTGTCTATTCGGATGACCTTCTT 1058  
Qy 564 CTCACACAGGACTACATCAATGCGCAGATGATCTGCTGATGACATGACCGACCAAGAA 623  
Db 1059 GATCATATCAACCAAGGACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1118  
Qy 624 TTTGATCGGCTTTTGGCTAGCGCATAGTTCTGACATTTGCTGGCTTTCTTCTCCTCT 683  
Db 1119 ACTCAATACTATTAGTGGTGAACCTTAATTTTGACTGCAACTACTTTGCTCCTCCCTTGG 1178  
Qy 684 TGGTCATTTTGGTGTCTATTTCACTGATGTTGTCAGGAGCTGATCAAGCGCAGGAGAAC 743  
Db 1179 TGATAGTGACATTTGCTATACCAAGATTTCCACACTCTGACCCATGGACTGCAACTG 1238  
Qy 744 TCATGAGACAGGACAAACACAGCGGAGCGAGCTGCTGCTGCGGACCATCTACTGTTGTG 803  
Db 1239 ACAGTGCCTTTAAGCAGAAAGCAGGAGCTTAACCAATTTCTCTACTCCGCGAGCGTGG 1298  
Qy 804 GC 805  
Db 1299 TC 1300

RESULT 13  
US-09-864-761-21546/c  
; Sequence 21546, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aesomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; PRIOR FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 21546  
; LENGTH: 348  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL009181.1  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.5  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.2  
; OTHER INFORMATION: EST\_HUMAN HIT: BE252309.1, EVALUE 2.10e-02  
; OTHER INFORMATION: NT HIT: g11526121, EVALUE 0.00e+00  
; OTHER INFORMATION: SWISSPROT HIT: P41143, EVALUE 6.00e-48  
; US-09-864-761-21546

Query Match 7.9%; Score 85.8; DB 10; Length 348;  
Best Local Similarity 55.6%; Pred. No. 7.8e-17;  
Matches 165; Conservative 0; Mismatches 132; Indels 0; Gaps 0;  
Qy 236 AACCAAGCGCTGGAGCTGTGCCACCACTATCTGTTGTAACCTGATGGTGGCGGACCTGCG 295  
Db 301 ACATAAGATGAAGAGCGGCCACCAACATCTACATCTTCAACCTGGCCTTAGCCGATGCGC 242  
Qy 296 TTTATGTCTATTGCCCTTCTCTCATCATCACTACTACTACTAGATGACAGGTGGCCCTTCG 355  
Db 241 TGGCCACCAAGCGCTGCTGCTTTCAGAGTGCCAAAGTACCTGATGGAGAGCGTGGCCCTTCG 182  
Qy 356 GGGAGCTGCTGCAAGCTGGTGCACCTTCCGTTCATATCAACCTTTACGGCAGCATCC 415  
Db 181 GCGAGCTGCTGCAAGCTGTGCTCTCCATCGACTACTCAATATGTTTACAGCATCT 122  
Qy 416 TGCTGCTGACCTGCTCTGTGTCACCAAGTTTCTAGTGTGTGTCACCACTGTTGTCG 475  
Db 121 TCACGCTCACCATGATGATGTTTTCACCGCTACATCGCTGTCTGCCACCTCTCAAGGCC 62  
Qy 476 TGCCTTACCGAGCCCGAGGAGCATGCTGGTGGCGACCAAGCTGATCAACATCTGTATCTGGTCTCG 532  
Db 61 TGGACTTCCGACGCTTCCCAAGCCCAAGCTGATCAACATCTGTATCTGGTCTCGTGG 5  
RESULT 14  
US-09-823-114-7  
; Sequence 7, Application US/09823114  
; Patent No. US20020061554A1  
; GENERAL INFORMATION:  
; APPLICANT: EVANS, CHRISTOPHER J.  
; APPLICANT: KEITH, DUANE E.  
; TITLE OF INVENTION: OPIOID RECEPTOR GENES

NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FORSTER  
STREET: 2000 PENNSYLVANIA AVENUE, NW, Suite 5500  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1888  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/823,114  
FILING DATE: 29-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/148,351  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 22000-20526.22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030 MRSNFOERSWSH  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1829 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 29..1144  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-823-114-7

Query Match 7.7%; Score 83.6; DB 10; Length 1829;  
Best Local Similarity 47.7%; Pred. No. 8.3e-16;  
Matches 275; Conservative 0; Mismatches 299; Indels 2; Gaps 1;  
155 TCTACCTCTCCCTGGCTACAGTATCATCTTTATCCTAGGCTGCCACTAAATGGCACTG 214  
177 TCGCCATACCGCGCTCTACTCGGCTGTGCGCAGTGGGCTTCTGGGCACTGTCTCG 236  
215 TCTTGTGGCACTCTCTGGGGCCAAACCAAGCGTGGAGCTGTGCCACCACCTATCTCGTGA 274  
237 TCATGTTGGCATCGTCCGGTACACCAAAATTGAAGACCGCCACCAACATCTACATCTTCA 296  
275 ACCTGATGTGGCGGACCTGCTTTATGCTATTGCTATTCCTTCATCATCACTACTAC 334  
297 ATCTGCTTTGGCTGATGCGTGGCCACAGCAGCTGCGCTTCCAGAGCGGCAAGTACT 356  
335 TAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGAAGCTGTGCACTTCTCTTCTATA 394  
357 TGATGGAACAGTGGCGCTTTGGGAGCTGCTGTGCAAGCTGTGCTCTCATTGACTACT 416  
395 TCACCTTTACGGCAGCATCTCTGCTGCTGACCTGCACTCTCTGTGACCACTAGTCTAGGTG 454  
417 ACAACATGTTCACTAGTACTCTTCCACCTCACCATGATGAGCGTGGACCGCTACATTCGTG 476  
455 TGTGCCACCACTGTCTTCTGCTGCCCTACCGGACCGCAGGATGCTGGCTGGGACCA 514  
477 TCTGCCATCTGTCAAGCCCTGGACTTCGGGACACAGCCAAAGCCCAAGCTGATCAATA 536  
515 GCACCACTGGGGCCCTGGTGGTCTCTCAGCTGCTGCCACACTGGCTTCTCCACCAAGG 574  
537 TATGCATCTGGGCTTGGCTTCAAGTGTCTGGGTCTCCCATCATGGTTCATGGCAGTACCC 596  
575 ACTACATCAATGCCAGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATG 634

Db 597 AACCCCGGATGCTGCAGTGGTATGCTGCTCCAGTTCGCCAGTCCAGCTGGTACTGGG 656  
QY 635 TTTTTCCTACGGCAGTATGTCGACATTTGTCGGCTT--TCITTCCTCTCTTGGTCAATTT 692  
Db 657 ACACCTGTGACCAAGATCTCGGTGTTCTCTCTTTGGCCTTCGTGGTGGCGATCCTCATCA 716  
QY 693 TGGTGTGCTATTACATGATGGTGCAGGAGCTGATCA 728  
Db 717 CGGTGTGCTATGCCCTCATGCTACTCGGCTGGCA 752  
RESULT 15  
US-09-214-904-3  
Sequence 3, Application US/09214904  
Patent No. US20010047519A1  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: TRANSGENIC ANIMAL IN WHICH THE EXPRESSION  
TITLE OF INVENTION: OF OPIATE RECEPTORS IS MODIFIED  
NUMBER OF SEQUENCES: 6  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/214,904  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR97/01282  
FILING DATE:  
APPLICATION NUMBER: FR 96.08810  
FILING DATE: 15-JUL-1996  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2218 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 58..1173  
US-09-214-904-3

Query Match 7.7%; Score 83.6; DB 10; Length 2218;  
Best Local Similarity 47.7%; Pred. No. 9.1e-16;  
Matches 275; Conservative 0; Mismatches 299; Indels 2; Gaps 1;  
QY 155 TCTACCTCTCCCTGGCTACAGTATCATCTTTATCCTAGGCTGCCACTAAATGGCACTG 214  
Db 206 TCGCCATACCGCGCTCTACTCGGCTGTGCGCAGTGGGCTTCTGGGCAAGTGTCTCG 265  
QY 215 TCTTGTGGCACTCTCTGGGGCCAAACCAAGCGTGGAGCTGTGCCACCACCTATCTGGTGA 274  
Db 266 TCATGTTGGCATCGTCCGGTACACCAAAATTGAAGACCGCCACCAACATCTACATCTTCA 325  
QY 275 ACCTGATGTGGCGGACCTGCTTTATGCTATTGCTATTCCTTCATCATCACTACTAC 334  
Db 326 ATCTGCTTTGGCTGATGCGCTGGCCACAGCAGCTGCGCTTCCAGAGCGGCAAGTACT 385  
QY 335 TAGATGACAGGTGGCCCTTCGGGGAGCTGCTCTGCAAGCTGTGCACTTCTCTTCTATA 394  
Db 386 TGATGGAACAGTGGCGCTTTGGGAGCTGCTGTGCAAGCTGTGCTCTCATTGACTACT 445  
QY 395 TCACCTTTACGGCAGCATCTCTGCTGCTGACCTGCACTCTCTGTGACCACTAGTCTAGGTG 454  
Db 446 ACAACATGTTCACTAGTACTCTTCCACCTCACCATGATGAGCGTGGACCGCTACATTCGTG 505  
QY 455 TGTGCCACCACTGTGTTGCTGCCCTACCGGACCGCAGGATGCTGGCTGGGACCA 514  
Db 506 TCTGCCATCTGTCAAGCCCTGGACTTCGGGACACAGCCAAAGCCCAAGCTGATCAATA 565



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 26, 2003, 21:20:18 ; Search time 1815 Seconds  
(without alignments)  
9654.830 Million cell updates/sec

Title: US-10-023-586b-1  
Perfect score: 1082  
Sequence: 1 atgtgttcatttcttcc.....ccagggttgacacatctgg 1082

oring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_nam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	200.6	18.5	2542	11 AK017378	AK017378 Mus muscu
2	200.6	18.5	3001	11 AK005013	AK005013 Mus muscu
3	141.2	13.0	641	14 BQ396255	BQ396255 NISC_ngl9
4	140.2	13.0	638	9 AL675845	AL675845 AL675845
5	128.2	11.8	899	14 BQ959110	BQ959110 AGENCOURT
6	123.2	11.4	623	14 BQ038875	BQ038875 pgnic.pk0

7	121.6	11.2	851	13 BI833118	BI833118 603090834
8	120.6	11.1	928	9 AL561864	AL561864 AL561864
9	120.4	11.1	613	9 AL657842	AL657842 AL657842
10	120.2	11.1	1063	14 BM918491	BM918491 AGENCOURT
11	119.8	11.1	408	13 BI401676	BI401676 MI-P-CPD-
12	119.6	11.1	1101	17 CNS0532S	AL18925 Tetraodon
13	113.8	10.5	955	9 AL521440	AL521440 AL521440
14	109.8	10.1	898	13 BI819396	BI819396 603034571
15	106.2	9.8	884	9 AL525059	AL525059 AL525059
16	106	9.8	553	13 BM031311	BM031311 496677 MA
17	103	9.5	606	17 AZ953874	AZ953874 2M02191L7
18	102.8	9.5	934	13 BI768397	BI768397 603053844
19	99.4	9.2	932	17 CNS025IR	AL211932 Tetraodon
20	97.8	9.0	484	14 BQ551383	BQ551383 H4008H11-
21	97.6	9.0	936	9 AL520218	AL520218 AL520218
22	96.6	8.9	1118	14 BM918898	BM918898 AGENCOURT
23	94.6	8.7	835	9 AL525098	AL525098 AL525098
24	94.2	8.7	663	13 BM426517	BM426517 pgf2n.pk0
25	91.6	8.5	467	10 AW323925	AW323925 uq41d11.y
26	91.4	8.4	720	13 BI753905	BI753905 603022907
27	90.4	8.4	491	12 BG712193	BG712193 pg11n.pk0
28	89.4	8.3	801	9 AL549441	AL549441 AL549441
29	89.2	8.2	549	12 BE938149	BE938149 MR3-TN004
30	89	8.2	764	12 BF345996	BF345996 602018717
31	88.8	8.2	609	9 AL588350	AL588350 AL588350
32	88.8	8.2	801	13 BG924078	BG924078 602823635
33	88.6	8.2	695	9 AJ450320	AJ450320 AJ450320
34	88.6	8.2	719	9 AJ455780	AJ455780 AJ455780
35	88.4	8.2	618	12 BF346620	BF346620 602021219
36	88.4	8.2	663	12 BG820137	BG820137 602782353
37	88.4	8.2	691	13 BG912683	BG912683 602807060
38	88.4	8.2	779	12 BG820495	BG820495 602782813
39	88.4	8.2	788	13 BI916752	BI916752 603178755
40	88.4	8.2	1083	14 BM805382	BM805382 AGENCOURT
41	88.4	8.2	1336	14 BM906977	BM906977 AGENCOURT
42	87.4	8.1	961	14 BQ056144	BQ056144 AGENCOURT
43	87.4	8.1	1139	14 BM922349	BM922349 AGENCOURT
44	86.8	8.0	678	10 BE314834	BE314834 601147561
45	86.2	8.0	1064	13 BI517798	BI517798 603042051

ALIGNMENTS

RESULT 1	AK017378	2542 bp	mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430432J15	linear	HTC 19-JAN-2002
LOCUS	AK017378		coupled 2, full insert sequence.		
DEFINITION	AK017378		HTC; CAP trapper.		
ACCESSION	AK017378		GI:12856588		
VERSION	AK017378				
KEYWORDS	AK017378		HTC; CAP trapper.		
SOURCE	AK017378		Mus musculus (strain:C57BL/6J) 6 days neonate head cDNA to mRNA, clone:5430432J15.		
ORGANISM	AK017378		Mus musculus		
REFERENCE	AK017378		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	AK017378		1		
TITLE	AK017378		Carninci, P. and Hayashizaki, Y.		
JOURNAL	AK017378		High-efficiency full-length cDNA cloning		
MEDLINE	AK017378		Meth. Enzymol. 303, 19-44 (1999)		
PUBMED	AK017378		99279253		
REFERENCE	AK017378		10349636		
AUTHORS	AK017378		2		
TITLE	AK017378		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	AK017378		Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
MEDLINE	AK017378		Genome Res. 10 (10), 1617-1630 (2000)		
PUBMED	AK017378		20499374		
	AK017378		11042159		

## REFERENCE AUTHORS

3 Shibata, K., Itoh, M., Alzawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kikunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system - 384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

	FEATURES	SOURCE
1	Age	1980 US Census
2	Sex	"
3	Educational attainment	"
4	Marital status	"
5	Occupation	"
6	Religious affiliation	"
7	Political party	"
8	Income	"
9	Home ownership	"
10	Health insurance	"
11	Life expectancy	1980 US Census
12	Mortality rate	"
13	Fertility rate	"
14	GDP per capita	World Bank
15	Population growth	"
16	Urbanization	"
17	Government expenditure	"
18	Private consumption	"
19	Investment	"
20	Exports	"
21	Imports	"
22	Balance of payments	"
23	Foreign debt	"
24	Official development assistance	"
25	Net capital flows	"
26	Current account balance	"
27	Trade balance	"
28	Export composition	"
29	Import composition	"
30	Export market share	"
31	Import market share	"
32	Export diversification	"
33	Import diversification	"
34	Export concentration	"
35	Import concentration	"
36	Export elasticity	"
37	Import elasticity	"
38	Export growth	"
39	Import growth	"
40	Export stability	"
41	Import stability	"
42	Export volatility	"
43	Import volatility	"
44	Export competitiveness	"
45	Import competitiveness	"
46	Export performance	"
47	Import performance	"
48	Export quality	"
49	Import quality	"
50	Export innovation	"
51	Import innovation	"
52	Export technology	"
53	Import technology	"
54	Export infrastructure	"
55	Import infrastructure	"
56	Export logistics	"
57	Import logistics	"
58	Export trade policy	"
59	Import trade policy	"
60	Export trade agreement	"
61	Import trade agreement	"
62	Export trade facilitation	"
63	Import trade facilitation	"
64	Export trade liberalization	"
65	Import trade liberalization	"
66	Export trade integration	"
67	Import trade integration	"
68	Export trade cooperation	"
69	Import trade cooperation	"
70	Export trade partnership	"
71	Import trade partnership	"
72	Export trade alliance	"
73	Import trade alliance	"
74	Export trade association	"
75	Import trade association	"
76	Export trade union	"
77	Import trade union	"
78	Export trade network	"
79	Import trade network	"
80	Export trade system	"
81	Import trade system	"
82	Export trade framework	"
83	Import trade framework	"
84	Export trade environment	"
85	Import trade environment	"
86	Export trade culture	"
87	Import trade culture	"
88	Export trade values	"
89	Import trade values	"
90	Export trade norms	"
91	Import trade norms	"
92	Export trade standards	"
93	Import trade standards	"
94	Export trade practices	"
95	Import trade practices	"
96	Export trade procedures	"
97	Import trade procedures	"
98	Export trade rules	"
99	Import trade rules	"
100	Export trade regulations	"
101	Import trade regulations	"
102	Export trade laws	"
103	Import trade laws	"
104	Export trade decrees	"
105	Import trade decrees	"
106	Export trade orders	"
107	Import trade orders	"
108	Export trade resolutions	"
109	Import trade resolutions	"
110	Export trade decisions	"
111	Import trade decisions	"
112	Export trade actions	"
113	Import trade actions	"
114	Export trade measures	"
115	Import trade measures	"
116	Export trade initiatives	"
117	Import trade initiatives	"
118	Export trade projects	"
119	Import trade projects	"
120	Export trade programs	"
121	Import trade programs	"
122	Export trade schemes	"
123	Import trade schemes	"
124	Export trade plans	"
125	Import trade plans	"
126	Export trade strategies	"
127	Import trade strategies	"
128	Export trade policies	"
129	Import trade policies	"
130	Export trade approaches	"
131	Import trade approaches	"
132	Export trade methods	"
133	Import trade methods	"
134	Export trade techniques	"
135	Import trade techniques	"
136	Export trade tools	"
137	Import trade tools	"
138	Export trade equipment	"
139	Import trade equipment	"
140	Export trade facilities	"
141	Import trade facilities	"
142	Export trade services	"
143	Import trade services	"
144	Export trade support	"
145	Import trade support	"
146	Export trade assistance	"
147	Import trade assistance	"
148	Export trade aid	"
149	Import trade aid	"
150	Export trade help	"
151	Import trade help	"
152	Export trade guidance	"
153	Import trade guidance	"
154	Export trade advice	"
155	Import trade advice	"
156		

GAGAGAGAGATCTCGAGTTTAATTAAATTAATCCCCCCCCCCCCC 3'. cDNA was cleaved  
with BamHI and XhoI. Vector: a modified pBluescript KS(+) after  
bulk excision from Lambda FLC 1. Cloning sites, 5' end: SalI; 3'  
end: BamHI. Host: DH10B.

	Location/Qualifiers
1..2542	
/organism="Mus musculus"	
/strain="C57BL/6J"	
/db_xref="PANTOM_DB:5430432J15"	
/db_xref="MGI:1897121"	
/db_xref="taxon:10090"	
/clone="5430432J15"	
/tissue_type="head"	

[illegible]



```
QY 604 GACATGACCAGCAAGAGAAATTTTGATCGCTTTTTCCTACGGCATATGTTCTG--ACAT 661
|||||  || |||||  |||||  || |||||  || |||||  || |||||  || |||||  || |||||
Db 1036 GACACCTCGCCGCGAGAGCTCTTTAGCCATTTTGTGGCTTACAGCTCCGCTCATGCTGGGT 1095

QY 662 TGTCTGGCTTCTTCTCCCTTCCTGGTCAITTTGGTGTGCTATTTCATGATGCTCAGGAGC 721
|||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||
Db 1096 CTGCTTTTGTGTGCGCTTTCCGTAATCTGCTGTGTTACGTGTGTTATGGCCAGCGGG 1155

QY 722 CTGATCAAGCCAGGAGAGAACCTTCATGAGGACAGGCAACACAGCCCGAGGCGGTCCATC 781
|||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||
Db 1156 CTGCTCAACCGGCTTATGGGACCACAGGAGGTCTGCCCTGGGCGCAAGCTGTGTG 1215

QY 782 CGGACCATCTACTGTGTGTGGCTCTTCACCCCTGTGTTGTGCGCTTCCATATCACT 841
|||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||
Db 1216 CGCACCATCTGCTGTGTTACTGCGCTCTTCGCGCTCTGCTCTTCTGCTTTCACCTCACG 1275

842 CGCTCTCTTACCTACCATCTCTCTTCTGCTTCTTCAGGACTGCCAGCTCTTGATGGCA 901
|||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||
Db 1276 CGCACCTCTCTACTACTCTCTTCCGAT--CACTTGACCTCAGCTGSCCAGCCCAACAGTTGCTTGACCGC 1332

QY 902 GCCAGTGTGGCTTACAGATATGAGGCGCTCTGCTGAGTGTGAGCAGCTGCCCTCAACCCCA 961
|||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||
Db 1333 ATCAACATGCGATATAAGATCACCGCGCTGSCCAGCGCCCAACAGTTGCTTGACCGC 1392

QY 962 GTCTGTACTTCTTCTTCAAGG 982
|||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||  || |||||
Db 1393 GTACTCTACTTCTCTGGCAGGG 1413

RESULT 2
AK005013 3001 bp mRNA linear HTC 19-JAN-2002
LOCUS Mus musculus adult male liver cDNA, RIKEN full-length enriched
DEFINITION library, clone:1300015C04:purinergic receptor p2y, G-protein
coupled 2, full insert sequence.
AK005013
VERSION AK005013.1 GI:12836638
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus
clone:1300015C04.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
PUBMED
AUTHORS
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
PUBMED
REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
PUBMED
REFERENCE
AUTHORS
```

```
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalow, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G.,
Quackenbush, J., Schriml, L. M., Staubli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, O., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
Toyo-Oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L.,
Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, J., Kohsaki, S.,
and Hayashizaki, Y.
Nature 409 (6821), 685-690 (2001)
21085660
11217851
PUBMED
REFERENCE
5 (bases 1 to 3001)
AUTHORS
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGAGAGAGCGGCCGCACTCGAGTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. Second
strand cDNA was prepared with the primer adapter of sequence[5'
GAGAGAGAGAGGATCAAGAGCTCAATTAATTAATTAACCCCCCCCCC 3']. cDNA was
cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end:
XhoI. Host: SOUR.
Location/Qualifiers
1. .3001
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:1300015C04"
/db_xref="MGI:1895057"
/db_xref="taxon:10090"
/clone="1300015C04"
/sex="male"
/tissue_type="liver"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. .3001
/gene="P2ry2"
gene
```

```
CDS
477..1598
/ gene="P2iy2"
/ note="data source:MGD, source key:MG1:105107,
evidence:ISS
purinergic receptor P2Y, G-protein coupled 2
putative"
/codon_start=1
/protein_id="BAB23746.1"
/db_xref="GI:12836639"
/db_xref="MGD:MG1:105107"
/translation="MAADLEPNSTINGTWEGDELGYKCFNEDFKVLLPVSGVVC
VLGLNVALYITFLCKATKNASTMYEHLVASDSLYAASLELLYIYARGDHPFS
TVLCKLVFLYTLNCSILFLTCISVHRCLGVLRLPLHSURWGRYARVAAVWVL
VLACQAVLYFVTSVRGTRITCHDSRELFSEFVAYSSVGLGLFAVPFVILVY
VLMARLLPAYGTGGLPRKRSVRTIALVLAFALCFPLFHLTRLYSFRSLDL
KTLNLINMAYKTRPLANSCLDPLVFLAGORLVRFAORDAKPTEPTSPQARR
KGLHRPNRTVRKDLVSSDDSRTESTPAGSETKDIRL"
polyA_signal
2982..2987
/ gene="P2iy2"
/ note="putative"
polyA_site
3001
/ gene="P2iy2"
/ note="putative"
BASE COUNT      630 a      861 c      834 g      676 t
IN
Query Match      18.5%; Score 200.6; DB 11; Length 3001;
Best Local Similarity 54.5%; Pred. No. 1.2e-43;
Matches 469; Conservative 0; Mismatches 384; Indels 8; Gaps 3;
Qy 127 TGCAGATTCAGAGAGTACAAGCAAGTCTACCTCCCTGGCGCTACAGTATCATCTTT 186
Db 549 TGTGCTTTCAACAGAGACTTCAAGTAGCTGTGTGGCCGTGTCTATGGCGTGTGTC 608
Qy 187 ATCTAGGCTGCCACCAATTAAGTGCCTCTGTGGCACTCTCTGGGCAACCAAGCGC 246
Db 609 GTGCTGGGTTGCTGGAAGCTGTGGCTCTCTATATCTTCTATGCGGCTCAAAACC 668
Qy 247 TGGAGCTGTGCCACCACTATCTTGTGTAACCTGTATGTTGGCCGACCTGCTTTATGCTA 306
Db 669 TGGAAAGCTTCCACCACTATCTTGTGTAACCTGTGCGAGTTTGGGCTCTCTACGACG 728
Qy 307 T---TGGCCTTCTTCATCATCACTACTACTAGATGACAGTGGCCCTTCGGGAGCTG 363
Db 729 TCCCTGGCGCTGTGGTTTATCTAGCCCGGGTGACCACTGCGCATTTAGCAGCGTG 788
Qy 364 CTCTGCAAGCTGGTGCATCTCTCTTATATCAACCTTTTACGGCAGCATCTCTGCTGCTG 423
Db 789 CTCTGCAAGCTGGTGGCTTCTCTTACACCAACCTCTACTGCAGCATCTCTTCTCCTC 848
Qy 424 ACCTGCATCTGTGCACCACTCTAGGTGTGTGCCACCACTGTGTGCTGCCCTAC 483
Db 849 ACCTGCATCAGCTGCACCGGTGCTGGAGTCTCTGGCCCTCTGCACCTCCCTGCGTTG 908
Qy 484 CGGACCCGACGATGCTGCTGGCAGCAGCACCACCTGGCCCTGGTGGTCTCCAG 543
Db 909 GCGCGGCCCTTATAGCCCGGGTGCTGCGGTGTGTGGGTGCTGGTGGTGGCTGCG 968
Qy 544 CTGCTGCCCACTGCGCTCTCCACACCGACTATACATCAATAGCCAGATGATCTGTTAT 603
Db 969 CAGGACCCGCTCTACTCTGTCACCAACAGCGTGGCGGGAACCCGGATCCTTGCCAT 1028
Qy 604 GACATGACCAAGCAAGAAATTTTATGCGGCTTTTTCCTACGGCATAGTTCTG--ACAT 661
Db 1029 GACACTCGGCCGAGAGCTCTTTAGCCATTTTCTGGCTTACAGCTCGCTCATCTGGGT 1088
Qy 662 TGTCTGGCTTTCTTCCCTCTCTGTCATTTTGTGTGCTATTACATCATGCTCAGGAGC 721
Db 1089 CTGCTTTTGTGTGGCCCTTTTTCGTAATCTGCTGTGTTACGTGCTTATGCCAGCGCG 1148
Qy 722 CTGATCAAGCCAGAGAGAACTCATAGGACAGGCAACACAGCCGACAGGCTCCATC 781
Db 1149 CTGCTCAACCGCGCTTATGGGACACACAGGAGTCTGCTCGGGCCACGCGAAGTCTGTG 1208

FEATURES
Source
1..641
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="IMAGE:5383884"
/tissue_type="neural"
/dev_stage="embryo, stages 14-19"
/lab_host="DH10B (phage-resistant)"
/notes="vector: pCMV-SPORT6.cdb; Site_1: NotI; Site_2:
EcoRV; Cloned unidirectionally. Primer: Oligo dT. Average
insert size 2.1 kb. Constructed by Invitrogen. Note: This
is a Xenopus Gene Collection (XGC) library."

BASE COUNT      132 a      184 c      153 g      172 t
ORIGIN
Query Match      13.0%; Score 141.2; DB 14; Length 641;
Best Local Similarity 55.5%; Pred. No. 1.2e-27;
Matches 294; Conservative 0; Mismatches 233; Indels 3; Gaps 1;
Qy 127 TGCAGATTCAGAGAGTACAAGCAAGTCTACCTCCCTGGCGCTACAGTATCATCTTT 186
Db 106 TGTAAATTCGACGAGGATTTTCAAGTAGCTCTCTCCCGCTGCTGACGATCTGTGTC 165
Qy 187 ATCTAGGCTGCCACATAAATGGCACTGTCTGTGGCACTCTCTGGGCAACCAAGCGC 246
```

```

Db 166 TGGTGGGGCTGATCTGAACATCTGGGCCCTGTATACATCTTCTCTGTTCCGGATCAAGCCC 225
QY 247 TGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCTA 306
Db 226 TGGAAAGCCCTCCACCACTACATGTTCAACCTGGCCATTTCCGACATGATGTACGTCATC 285
QY 307 T---TGGCCCTTCTCTATCATCACTACTACTAGATGACAGTGGGCCCTTCGGGGAGCTG 363
Db 286 TCCTCCCGCTGCTGTCTATTACTACTCGCAGGGGACAACTGGCCGTTGCGCGTGGCC 345
QY 364 CTCTGCAAGCTGGTGCACATCTCTGTCTATATCAACCTTTACGGCAGCATCTCTGCTG 423
Db 346 TTGTGCAAGATGCTCAAGTCTCTTTCTACACCAACATGTACTGTAGCATCTCTCTCTTG 405
QY 424 ACCTGCATCTCTGTGCACAGTTCTCTAGTGTGTGCCACCCACTGTGTCTGCTGCCCTAC 483
Db 406 CTGTGTCATCAGCATCACCGCTCTCTCGGATCTGTACCCATGAATCACTCGGTTGG 465
QY 484 CGGACCCGAGGATGCTGTGGTGGGACACACACCTGGGCCCTGTGTGTCCTCCAG 543
Db 466 CTGAAGGTGGGAATGCTCGGATTATTTCCGTGGTCTGCGTTTATCGTTTCCGCTGT 525
QY 544 CTGCTGCCACACTGCTCTCCACAGGACTACATCAATGCCAGATGATCTGTGAT 603
Db 526 CAGTCCCTATATTGTACTTGTATACCAACAGTTCCCAATGGGGACAGCACCCTGCCAT 585
QY 604 GACATGACCAAGCAAGAAATTTGATCGCTTTTTCGCTACGGCATAGT 653
Db 586 GACACTTCCAGCGTGATCTATTGTGACAACTTTGTGCTCTACAGCAGT 635

RESULT 4
AL675845
LOCUS
DEFINITION AL675845 XGC-gastrula Silurana tropicalis cDNA clone Tgas05ln19 5',
            mRNA sequence.
ACCESSION AL675845
VERSION AL675845.1 GI:19532219
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
            Xenopodinae; Silurana.
            1 (bases 1 to 638)
            Taylor R., Ashurst, J. L., Croning, M. D. R., Zorn, A. M. and Rogers, J.
            Sanger xenopus tropicalis EST project 2002
            Unpublished (2001)
            Contact: Taylor R
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: Tgas05ln19.picSP6
            Sequencing primer: PICSP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
            Location/Qualifiers
                1..638
                /organism="Silurana tropicalis"
                /db_xref="taxon:8364"
                /clone="Tgas05ln19"
                /clone_lib="XGC-gastrula"
                /dev_stage="gastrula (stages 10 5-13 mixed)"
                /lab_host="Escherichia coli XL1-blue"
                /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                was oligo df primed from 5ug of poly A+ RNA from stages
                10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
                into pCS107 with EcoRI at the 5' end and NotI at the 3'
                end."
BASE COUNT 134 a 180 c 144 g 179 t 1 others
ORIGIN

```

```

Query Match 13.0%; Score 140.2; DB 9; Length 638;
Best Local Similarity 55.3%; Pred. No. 2.3e-27;
Matches 293; Conservative 0; Mismatches 234; Indels 3; Gaps 1;

QY 127 TGCAGTTCTCAGAGATACAAAGTCTACTCTCCCTGGCCTACATATCATCTTT 186
Db 102 TGTAAATTCAGCAGGAGATTTCAGAGTACGTCTCTCCCGTGTGTCAGGCATCGTGTTC 161
QY 187 ATCTTAGGCTGCCATTAATGGCACTGTCTTGTGGCACTCTCTGGGCCAAACCAAGCGC 246
Db 162 TGGTGGGGCTGATCTTGAACATTTGGCCCTGTATACATCTTCTCTGTTCCGGATCAAGCCC 221
QY 247 TGGAGCTGTGCCACCACTATCTGTGTGAACCTGATGGTGGCCGACCTGCTTTATGTGCTA 306
Db 222 TGGAAAGCCCTCCACCACTACATGTTCAACCTGGCCATTTCCGACATGATGTACGTCATC 281
QY 307 T---TGGCCCTTCTCTATCATCATTACTACTAGATGACAGTGGGCCCTTCGGGGAGCTG 363
Db 282 TCCTCCCGTGTGCTGTATTACTATCGCAGGGGACAACTGGCCGTTTCGGCGTGGCC 341
QY 364 CTCTGCAAGCTGGTGCACATCTCTGTGTCTATATCAACCTTTACGGCAGCATCTCTGCTGT 423
Db 342 TTGTGCAAGATCGTCAAGTCTCTTTCTACCAACATGTACTGTAGCATCTCTCTCTTG 401
QY 424 ACCTGCATCTCTGTGCACCACTCTCTAGTGTGTGCCACCCACTGTGTGCTGCCCTAC 483
Db 402 CTGTGTCATCAGCATCCACCGCTTCTCTCGGATCTGTACCCAAATGAATCACTCGGTTGG 461
QY 484 CGGACCCGAGGATGCTGTGGTGGGACACAGCACCACCTGGGCCCTGTGTGCTCTCCAG 543
Db 462 CTGAAGGTGGGAATGCTCGGATTATTTCCGTGGTCTCTGCGTTTATCGTTTCCGCTGT 521
QY 544 CTGCTGCCACACTGCTCTTCTCCACAGGACTACATCAATGCCAGATGATCTGTGAT 603
Db 522 CAGTCCCTATATTGTACTTGTATACCAACAGTTCCCAATGGGGACAGCACCCTGCCAT 581
QY 604 GACATGACCAAGCAAGAAATTTGATCGGCTTTTTCGCTACGGCATAGT 653
Db 582 GACACTTCCAGCGTGGATCTATTGTGACAACTNTGTGCTCTACAGCAGT 631

RESULT 5
BO959110
LOCUS
DEFINITION BO959110 899 bp mRNA linear EST 21-AUG-2002
            5', mRNA sequence.
ACCESSION BO959110
VERSION BO959110.1 GI:22374588
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 899)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs@mail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2651 row: e column: 10
            High quality sequence stop: 677.
            Location/Qualifiers
                1..899
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:6462033"
FEATURES
            source

```

```

/clone_lib="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site:1: ECORT; Site:2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into ECORI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
131 a 301 q 175 t
BASE COUNT 131 a 292 c 301 q 175 t

```

[illegible]

LOCUS	DEFINITION	623 bp	· mRNA	linear	EST 01-MAY-2002
BQ038875	gpmic.p0k010.1l13 normalized chicken lymphoid cDNA library				
	gallus cDNA clone gpmic.p0k010.1l13 5' similar to				
	sp134966.P2YR.CHICK P2Y PURINORECEPTOR 1 (ATP RECEPTOR) (P2Y1)				
	(PURINERGIC RECEPTOR) p1r1S37333 G protein-coupled receptor -				
	chicken emb1/CAA51716.1 (X73268) ATP receptor P2Y1 [Gallus gallus],				
	mRNA sequence.				

accession B0038875  
 version B0038875.2  
 keywords GI:20383637  
 source chicken.  
 organism Gallus gallus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

RESULT 7	
BI83118	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	





TITLE	JOURNAL	MEDLINE	COMMENT
1. <i>Effect of a 10% increase in the number of patients on the waiting time for a general practitioner consultation</i>	<i>Journal of the Royal Society of Medicine</i>		
2. <i>The effect of a 10% increase in the number of patients on the waiting time for a general practitioner consultation</i>	<i>Journal of the Royal Society of Medicine</i>		
3. <i>The effect of a 10% increase in the number of patients on the waiting time for a general practitioner consultation</i>	<i>Journal of the Royal Society of Medicine</i>		
4. <i>The effect of a 10% increase in the number of patients on the waiting time for a general practitioner consultation</i>	<i>Journal of the Royal Society of Medicine</i>		
5. <i>The effect of a 10% increase in the number of patients on the waiting time for a general practitioner consultation</i>	<i>Journal of the Royal Society of Medicine</i>		
6. <i>The effect of a 10% increase in the number of patients on the waiting time for a general practitioner consultation</i>	<i>Journal of the Royal Society of Medicine</i>		
7. <i>The effect of a 10% increase in the number of patients on the waiting time for a general practitioner consultation</i>	<i>Journal of the Royal Society of Medicine</i>		
8. <i>The effect of a 10% increase in the number of patients on the waiting time for a general practitioner consultation</i>	<i>Journal of the Royal Society of Medicine</i>		
9. <i>The effect of a 10% increase in the number of patients on the waiting time for a general practitioner consultation</i>	<i>Journal of the Royal Society of Medicine</i>		
10. <i>The effect of a 10% increase in the number of patients on the waiting time for a general practitioner consultation</i>	<i>Journal of the Royal Society of Medicine</i>		

Iowa State University  
201 Klidde Hall, Ames, IA 50011-3150, USA  
Tel: 5152944252  
Fax: 5152942401  
Email: cktugle@iastate.edu  
Oligo-dr track not found, Not I site shown in beginning of sequence  
is likely internal to the message. cDNA Library Preparation: M.B.  
Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab,  
University of Iowa clone distribution: clones will be available  
through Research Genetics ([www.resgen.com](http://www.resgen.com))  
Seq primer: M13 Forward  
POLYA=No.

Soares Lab, University of Iowa EST sequencing; M.B. Soares Lab,  
University of Iowa Clone Distribution; clones will be available  
through Research Genetics ([www.resgen.com](http://www.resgen.com))  
Seq primer: M13 Forward  
POLYA=No.

Location/Qualifiers
1. 408
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="MI-P-CP0-nvq-a-03-0-UI"
/clone_lib="MI-P-CP0"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: EcoRI; The MI-P-CP0 library is derived from uterus. For a detailed description of the library, from which this clone was derived, please visit our web site at <a href="http://pigst.genome.iasstate.edu/">http://pigst.genome.iasstate.edu/</a> . The procedure used to create this library has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)

Db	103	CTCTGCAAGCTGGTGGCGCTTCTTCTACACCAACCTGTACTGCAGCATCCTGTTCTC	44
QY	424	ACCTGCACTCTCTGTGCACCAAGTTCCTAGGTGT	455
Db	43	ACCTGCAATCAGCGTGCACCGCTGCGTGGCGGT	12

RESULT 12	CNS0532S/c	CNS0532S	1101 bp	DNA	linear	GSS 26-JUL-2000
LOCUS						
DEFINITION		Tetraodon nigroviridis genome survey sequence T7 end of clone 020M21 of library A from Tetraodon nigroviridis, genomic survey sequence.				
ACCESSION		AL318925				

VERSION AL318925.1 GI:9551809  
 KEYWORDS GSS: genome survey sequence.  
 SOURCE Tetraodon nigroviridis.  
 ORGANISM Tetraodon nigroviridis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
 Tetraodontidae; Tetraodon.  
 REFERENCE 1 (bases 1 to 1101)  
 AUTHORS Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
 Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,  
 Saurin, W. and Weissenbach, J.  
 TITLE Estimate of human gene number provided by genome-wide analysis  
 using Tetraodon nigroviridis DNA sequence  
 JOURNAL Nat. Genet. 25 (2), 235-238 (2000)  
 MEDLINE 20296633  
 PUBMED 10835645  
 REFERENCE 2 (bases 1 to 1101)  
 AUTHORS Crolius, H.R., Jaillon, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C.,  
 Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W.,  
 Bernot, A. and Weissenbach, J.  
 TITLE Characterization and repeat analysis of the compact genome of the  
 freshwater pufferfish Tetraodon nigroviridis  
 JOURNAL Genome Res. 10 (7), 939-949 (2000)  
 MEDLINE 20359837  
 PUBMED 10899143  
 REFERENCE 3 (bases 1 to 1101)  
 AUTHORS Genoscope.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-APR-2000)  
 COMMENT This sequence is a single read and was generated as part of a large  
 scale clone and sequencing project of the Tetraodon nigroviridis  
 genome. For more information, please take a look at  
 http://www.genoscope.cns.fr/tetraodon.  
 FEATURES  
 source  
 1. 1101  
 /organism="Tetraodon nigroviridis"  
 /db\_xref="taxon:99883"  
 /clone="020M21"  
 /clone\_lib="A"  
 /note="Genoscope sequence ID : COA020AG11C1-end : T7"  
 BASE COUNT 305 a 216 c 280 g 284 t 16 others  
 ORIGIN  
 Query Match 11.1%; Score 119.6; DB 17; Length 1101;  
 Best Local Similarity 53.3%; Pred. NO. 1.2e-21;  
 Matches 288; Conservative 0; Mismatches 247; Indels 5; Gaps 2;  
 QY 116 AACAGGCTCTGCAGCTCTCAGAGAAGTACAAAGTCTACCTCTCCCTGGGCTACA 175  
 DB 537 ATCAACTACAGTGCCTTTGACGAGGACTTCAATATACCTCTCTGCCGCTGTCTATG 478  
 176 GTATCATCTTTATCTAGCGCTGCCACTAAATGGCACTGTCTGTGGCACTCTCTGGGGCC 235  
 DU 477 GCCTGGCTTTGTCTTCGGCTTCATCTCAAGCCACAGCCCTGTGGCTCTCTCTGA---421  
 QY 236 AAACAAGCGTGGAGCTGTGCCACCACTATCTGGTGAACCTGATGTGGCCGACCTGC 295  
 DB 420 AGATCGGCGCTGGAACCCCAACACCGCTCTTCTTGTCCACCTCGCCCTCTCCGACTTC 361  
 QY 296 TTATGTGCTATTGGCTTCTCTCATCATCACTACTACTAGATGACAGGTG--GCCCTT 353  
 DB 360 TCTACATTTGTGCGTGGCCACCTCTATCTATATAGCCCAACCCAGCCGCACTGGCCCTT 301  
 QY 354 CGGGAGCTGTCTGCAAGCTGGTGCATCTCTCTGTCTATATCAACCTTTACGGCAGCAT 413  
 DB 300 CGGCGTGCAGCTGCAAAATTGGCGCTCTCTCTCTCTACGCTCAACCTCTACTGCAGCAT 241  
 QY 414 CTTGCTGTGACCTGCATCTCTGTGCACAGCTTCTTAGTGTGTGCCACCCACTGTGTC 473  
 DB 240 CTTCTTCCTACCTGCATCGCGTGTACCGTTACGTGGCATCTGCCACCCGATTAAGGC 181  
 QY 474 GCTGCCCTACCGGACCCGAGCATGCGCTGGCTGGCGCCACGACCACTGCGCCCTGTGT 533

Db 180 GCTGAATCTGCTGAANCCCGGANATTCCTACCTGTGTGGCTGGTGTGGGCTGTGTGT 121  
 QY 534 GGTCTCTCCAGCTGCTGCCACACACTGGCTCTCCACACGGGACTCATCAATGGCCAGAT 593  
 Db 120 CATAGTCTGTCTGGTGGCTTAACCTCATCTTTGTACCACCTCCAGGCGGGAATATCNC 61  
 QY 594 CATCTGGTATGATGACAGCAGCAAGAAATTTTGTATCGCTTTTGGCTAGGCATAGT 653  
 Db 60 CCTGTGCCATGATACACACNCCGCGAGTATTCTTGCATATATGATAATTAAGTGCNNCT 1  
 RESULT 13  
 LOCUS AL521440 955 bp mRNA linear EST 13-FEB-2001  
 DEFINITION AL521440 LTI\_NFL004\_NBC2 Homo sapiens cDNA clone CS0DB001YJ16.5  
 prime, mRNA sequence.  
 ACCESSION AL521440  
 VERSION AL521440.1 GI:12784933  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 955)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.  
 Location/Qualifiers  
 1. 955  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DB001YJ16"  
 /clone\_lib="LTI\_NFL004\_NBC2"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
 vector. Library was normalized. Library was constructed  
 by Life Technologies. Contact : Feng Liang Life  
 Technologies, a division of Invitrogen 9800 Medical Center  
 Drive Rockville, Maryland 20850, USA Fax : (1) 301 610  
 8371 Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com"  
 BASE COUNT 160 a 400 c 159 g 209 t 27 others  
 ORIGIN  
 Query Match 10.5%; Score 113.8; DB 9; Length 955;  
 Best Local Similarity 49.7%; Pred. NO. 4.5e-20;  
 Matches 283; Conservative 20; Mismatches 260; Indels 6; Gaps 2;  
 QY 126 CTGCAGTCTTCAGAGAAGTACAAAGTCTACCTCTCCCTGGCGCTACAGTATCATCTT 185  
 DB 285 CTGTCTCTACCGGAGAACTTCAAGCACTGCTGTGCCACCTGTGTATTCGCGGTGCT 344  
 QY 186 TATCCTAGGGTGGCCATTAATGGCACTGTCTTGTGGCACTCTCTGGGGCAACCAAGG 245  
 DB 345 GCGGCTGGCGCTGGCGTGAACATCTGTCTATTACCCAGATCTGCAGTCCCGCGGCG 404  
 QY 246 CTGGAGCTGTCCACCTATCTGTGAACCTGATGTGGCCGACCTCTTATGTGTGCT 305  
 DB 405 CTTGACCCGCGCGCGGTGACACCTTAACCTTCTGTGGCTGACCTGCTATATGCGCTG 464  
 QY 306 AT---TGCCCTTCTTCATCATCACTACTACTAGATGACAGTGGCCCTTCGCGGAGCT 362  
 DB 465 CTCCTGCCCCCTGCTCATCTACAACTATGCCCAAGGTGATCACTGGCCCTTTGGCGANTT 524



```
QY 363 GCTCTGCAAGCTGGTCACTTCTCTCTTATATCAACCTTTACGGCAGCATCTGCTGCT 422
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 525 CGCCTGCMSCCTGGTCCGCTTCTCTCTATGCCAACCTGAACGGCAGCATCTCTTCT 584
QY 423 GACCTGCATCTCTGTCACCACTGTCCTAGGTGTGGCCACCACTGTGTGGTGGCCCTA 482
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 585 CACCTGTCATCTGTCACCGCTACCTGGGCATCTGCGACCGCTGGCACCTGGCACAA 644
QY 483 CC---GGACCCGAGCATGCTGGCTGGGCGACCACTGGGCGCTGTGTGCTCT 539
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 645 ACSTGSSSCSCSSCTSSCTSTAGTGTSTAGCCCTSTSCCTCCCTCCACAA 704
QY 540 CCAGCTGTGGCCACACTGGCTTCCACACAGGACTACATCAATGGCCAGCATGCTG 599
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 705 CCACCTCCCTCCACACCACTATTTBCTCCACACCACTMCAACCCCACTCTCTC 764
QY 600 GTATGACATGACCAACCAAGAAATTTTGATGGCTTTTGGCTAGCGGATAGTCTGAC 659
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 765 CTATCAACTACCCCTCCCTCCCTCCCTCCCTCCCTATCCCTATCCCTATCCCTCTCAC 824
QY 660 ATTGTGCTGCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 688
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 825 TCTCATYCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 853

RESULT 14
LOCUS B1819396
DEFINITION 603034571F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:517561 5',
mRNA sequence.
ACCESSION B1819396
VERSION B1819396.1 GI:15930946
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 898)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM1437 row: n column: 18
High quality sequence stop: 884.
Location/Qualifiers
1..898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:517561"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
149 a 303 c 255 g 191 t

BASE COUNT
ORIGIN
```

```
Query Match 10.1%; Score 109.8; DB 13; Length 898;
Best Local Similarity 53.0%; Pred. No. 5.5e-19;
Matches 329; Conservative 0; Mismatches 282; Indels 10; Gaps 4;

QY 126 CTGCGCAATGTTCTAGAGAAAGTACAAGCAAGTCTACCTCTCTCCCTGGCCCTACAGTATCATCTT 185,
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 242 CTGTGTCTTACCGGAGAACTTCAAGCAACTGCTGCTGCCACCTGTGTATTCGGCGGTGCT 301
QY 186 TATCCTAGGCTGCTCAATAATGGCACTGCTTGTGTGGCACTCTCTGGCCGCAACCAAGCG 245
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 302 GCGCGCTGGCTGCCCTGACATCTGTGTCACTTACCAGATCTGCAGCTCCCGCGGGC 361
QY 246 CTGAGCTGTGGCCACCACTATCTGTGAACCTGTGTGGCCGACCTGTCTTATGTGCT 305
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 362 CCGTACCGCGCAGCGCTGTACACCTTAAACCTTGTCTTGTGCTGACCTGATATGCGCTG 421
QY 306 AT---TGCCCTTCTCATCATCACTACTACTAGATGACAGCTGGCCCTTCGGGAGCT 362
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 422 CTCCTTGGCCCTGCTCATCTACAACCTATGCCAAGGTGATCACTGGCCCTTTGGCGACTT 481
QY 363 GCTCTCAAGCTGCTGCACCTTCTCTTATATCAACCTTTAGCGGACGATCTCTGCTGCT 422
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 482 CGCTGCGCGCTGTGCTCTCTCTATGCCAACCTGCACGGCAGCATCTCTTCTCT 541
QY 423 GACCTGCATCTCTGTGCACCACTTCTGTAGTGTGTGCCACCACTGTGTGGTGGCCCTA 482
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 542 CACCTGCATCAGCTTCCAGCGCTACCTGGGCATCTGCCACCCGCTGCCGCCCTGGCAAA 601
QY 483 CC---GGACCCGAGCATGCTGGCTGGGCGACCACTGGGCGCTGGGCGCTGGTGTCTC 538
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 602 ACTGGGGGCGCTTGGCGGCTGCTGCTGTGTGTAGCCGTGTGCTGGCCGTGACAA 661
QY 539 TCCAGCTGTGTGCCACACTGGCTTCTTCCACACGGAATACATCAATGGCCAGATGATCT 598
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 662 CCAGTGCCTTGGCCAGCCATCTTCTGCTGCCACAGCATCCAGCGTAAACCGCATCTCT 721
QY 599 GGTATGACATGACCAAGCAAGAAATTTGATCGGCTTTTTC-CCTACGGCAGTAGTTCTG 657
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 722 GCTATGACCT--CAAGCGCGCTGCTGGGCGACCACTATATGCCCTATGCGATGCGCTCTC 779
QY 658 ACATTGCTGTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 717
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 780 ACTGTCATCGGCTTCTGTGCTGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 718 GAGCCTGATCAAGCCAGAGGA 738
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 840 GTCCCGCAGGATGCGCGGAGA 860

RESULT 15
LOCUS AL525099
DEFINITION AL525099 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC0051M02 5
prime, mRNA sequence.
ACCESSION AL525099
VERSION AL525099.1 GI:12788592
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 884)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..884
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

